A chemo-enzymatic cascade for the production of enantiopure aminoalcohols

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Elia Calderini
aus Città di Castello / Italien

1. Referentin: Professorin Dr. Anett Schallmey
2. Referentin: Professorin Dr.-Ing. Antje Spieß
eingereicht am: 19.06.2019
mündliche Prüfung (Disputation) am: 05.09.2019

Druckjahr 2019

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch die Mentorin der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

Elia Calderini, Julia Wessel, Patrick Schrepfer, Philipp Süss, Rainer Wardenga and Anett Schallmey. Selective Ring-Opening of Di-substituted Epoxides Catalysed by Halohydrin Dehalogenases, ChemCatChem **2019**, *11* (8), 2099–2106. https://doi.org/10.1002/cctc.201900103.

Tagungsbeiträge

Elia Calderini, Julia Koopmeiners, Rainer Wardenga and Anett Schallmey. Selective ring opening of di-substituted epoxides catalysed by halohydrin dehalogenases. (Vortrag). 2017, Microbiology and Infection 2017 - 5th Joint Conference of the DGHM & VAAM, Würzburg, Germany.

Elia Calderini; Philipp Süss; Rainer Wardenga and Anett Schallmey. A chemo-enzymatic cascade for the production of enantiopure amino alcohols. (Vortrag). 2018, 8th International Cebitec Research Conference, Bielefeld, Germany.

Posterbeiträge

Elia Calderini, Julia Koopmeiners, Rainer Wardenga and Anett Schallmey. Selective ring opening of di-substituted epoxides catalysed by halohydrin dehalogenases. Biocat 2016 International Congress on Biocatalysis, Hamburg, Germany.

Elia Calderini; Julia Koopmeiners; Janine Mayer; Rainer Wardenga and Anett Schallmey. Rosetta-based protein engineering of HheG. Biotrans 2017, 13th International Symposium on Biocatalysis and Biotransformations. Budapest, Hungary.

Elia Calderini, Julia Koopmeiners, Rainer Wardenga and Anett Schallmey. Selective ring opening of di-substituted epoxides catalysed by halohydrin dehalogenases. SUMMER SCHOOL 2017 on Biocatalysis as a Key Enabling Technology. Siena, Italy

Elia Calderini; Philipp Süss; Rainer Wardenga and Anett Schallmey. A chemo-enzymatic cascade for the production of enantiopure amino alcohols. 2018, 8th International Cebitec Research Conference, Bielefeld, Germany

Elia Calderini; Philipp Süss; Rainer Wardenga and Anett Schallmey. A chemo-enzymatic cascade for the production of enantiopure amino alcohols. 2018, Gordon Research Conference (GRC) and Gordon Research Seminar (GRS) on Biocatalysis, Biddeford, ME, US.

Elia Calderini, Julia Wessel, Philipp Süss, Rainer Wardenga and Anett Schallmey. Halohydrin dehalogenase-catalyzed azidolysis of vicinally di-substituted epoxides. 2018, Novel Enzymes, Darmstadt, Germany.

Student work and external contributions

Section 3.2.2.1: The mutants suggested by the *in silico* mutagenesis were generated and characterized by Janine Mayer in her Master thesis project (2016/2017).

Section 3.1.2.2: Epoxide ring opening reactions of substrates **23-25** were carried out by Julia Koopmeiners.

ABSTRACT

The increasing demand for “greener” products in chemical and biochemical processes made enzymatic or chemo-enzymatic cascades a very popular approach. Combining different reactions in one pot without intermediate purification steps represents numerous advantages compared to the classic approach. In a classic organic synthesis, longer operation time is needed and lower yields are typically reached. In addition, the elimination of work up steps in between reactions significantly reduces the amount of waste improving the environmental impact of a given process. The use of chemo-enzymatic cascades is particularly suitable for pharmaceutical processes where very specific diastereoisomers need to be synthesised. The priority is typically given to selectivity over high volumetric productivity. Furthermore, combination of chemical and enzymatic catalysts allows to access the complete set of reactions available for organic synthesis. However, several challenges arise as often the reaction conditions are incompatible and require extensive optimization.

In this thesis, the establishment of a chemo-enzymatic cascade for the synthesis of enantiopure aliphatic amino or azido alcohols starting from achiral non-terminal alkenes is presented. Optically active amino alcohols are an important class of compounds that find applications mostly in pharmaceutical products, as well as agrochemicals and the cosmetic or food industry. The new synthetic route was assembled using either two enzymes and one chemical catalyst or two chemical catalysts and one enzyme to enable the synthesis of both opposite amino alcohol enantiomers starting from prochiral alkenes. This was achieved by using either an enzymatic or a chemical epoxidation in the first step to selectively synthesise opposite enantiomers. The second step involved an enzymatic epoxide ring opening to control the regioselectivity of the nucleophilic attack. Finally, the resulting azido alcohol was chemically reduced affording enantiopure (2*S,3R)*-2-amino-3-hydroxyhexane, (2*S,3R)*-2-amino-3-hydroxyheptane, (2*R,3S)*-2-amino-3-hydroxyhexane and (2*R,3S)*-2-amino-3-hydroxyheptane. During this project, suitable catalysts for each step have been selected and optimal conditions for each step could be identified. For the combination of the different steps in a cascade, fine tuning of the reaction conditions was required to afford sufficient isolated yields and enantiopurity. Additionally, the catalytic scope of HHDHs was further expanded by in depth characterization of the regio- and enantioselectivity of a set of 22 HHDHs, representing all phylogenetic subtypes, in the epoxide ring opening of racemic, vicinally di-substituted *trans*-epoxides.

Table of contents

1. Introduction

1.1 White biotechnology

The use of biotechnology dates backwards to 6000 BC,[1] long before the existence of enzymes was even discovered. The use of yeast in fermentations to make bread, wine and cheese or to brew beer can be considered a biotechnological method. The name 'enzyme' is derived from ancient Greek and means "in-yeast" (en= in, zyme= leaven).[2] To this day, food and beverages technology represents a major portion of the total industrial enzyme market.[3] Enzyme's global market value is set to grow about 7% per year to reach a value of approximatively $ 6.2 billion (market value per year) in 2020.[1] Currently, the application of enzymes in other industrial processes is limited compared to the many application of classic chemical reactions (Figure 1). So far, most of the enzymes that meet industrial criteria belong to the class of hydrolases (e.g. amylases, proteases, cellulases and lipases), although processes involving other enzyme classes can be found.[4] One reason for this may be that enzymes evolved in cellular environment where substrate concentrations are far from the typical industrial requirements. Therefore, not many native enzymes could outperform the available chemical catalysts. In the last couple of decades, with the development of sophisticated technologies for DNA manipulation more examples of industrial applications of enzymes are emerging. These technologies made possible to engineer proteins to match certain process specifications. Bornscheuer *el at.* described the development of enzymatic catalysts as waves.[5] The first wave was characterised by the use of homogenized animal tissues or plant extract known to perform a certain reaction. At this point, the responsible enzymes able to catalyse the reaction were unknown. Around the 1980s, the newly discovered molecular technologies allowed the isolation and recombinant production of enzymes of interest, it was the beginning of the second wave. A decade later, the establishment of advanced DNA manipulation techniques to introduce mutations and improve enzymes for a specific function dictated the beginning of the third wave.[5] This strategy allowed a more efficient improvement of enzymes and is now commonly addressed as "directed evolution".[6,7]

Figure 1. Fields where enzymes are applied.

Recently, the importance of the third wave was recognized with the Nobel Prize in Chemistry won by Frances Arnold, George Smith and Sir Gregory Paul Winter for their work on directed evolution of enzymes and for phage display of peptides and antibodies, respectively. According to Bornscheuer, we are currently on the edge of the fourth wave.[6] This new period will be characterized by the exponential development of computer technology. It is already evident that the progress made on algorithms for machine learning and big data analysis is huge. Therefore, in the close future it is likely that sequence-function relations, the tertiary structure of proteins and even correct mutations to improve a certain characteristic could be predicted with extremely high accuracy. Additionally, stringent laws on environment preservation may be indirectly responsible for the recent boost in the development of enzyme-based processes. Enzymes generally work at milder reaction conditions using renewable materials (preferably) and often they allow the elimination of toxic components. All these characteristics make enzymes very suitable candidates for the Green Chemistry guidelines. However, normally intense enzyme and process engineering are required to develop an enzyme-catalysed industrial process.

1.1.1 Biocatalysis

Biocatalysis is generally defined as the use of free enzymes or whole cells in catalytic conversions. This kind of reactions can be further classified in two types: the first and oldest known type involves the use of growing cells and a carbon source. The carbon source is needed to ensure cellular growth as, in this case, the cellular metabolism is

exploited to produce proteins, primary or secondary metabolites or to convert a certain substrate. This type is also commonly referred to as fermentation. The second type involves the use of isolated enzymes, cell-free extract or resting cells for the conversion of desired substrates. In this case, catalyst production and substrate conversion are carried out in different stages. Fermentation was a known process for many centuries as mentioned previously, however, biocatalysis was not taken seriously until the 1980s where DNA technology developed exponentially. Additionally, stricter legislation for commercial drugs has indirectly influenced the development of biocatalysis. These laws required to test the activity of isomers of chiral drugs separately forcing chemists to look for alternative catalysts for asymmetric synthesis (e.g. enzymes). Thus, the use of enzymes in catalysis started to be more popular and the publication of Zaks and Klibanov in 1984 can be considered the first milestone.[8] They demonstrated that some enzymes, lipases in their case, could work well in organic solvents and at elevated temperatures. This represented a groundbreaking discovery as it was widely accepted that enzymes could only work in water. Hence, organic chemists suddenly had a whole range of new catalysts in their hands as an alternative to classic chemical catalysts. As mentioned earlier, directed evolution contributed heavily to the adoption of enzymes in industrial processes. Before this technology, it was necessary to adjust the conditions of a given process to the required optimal conditions of a given biocatalyst. With the advent of DNA manipulation techniques and high throughput screening, it was possible to do the exact opposite. Nowadays, industrial enzymes are almost entirely evolved in vitro and further formulated to fit the requirements of the process. This significantly lowers the investment for their integration in pre-existing technologies as enzymatic reactions can be run in multipurpose batch reactors. Currently, enzymes are mostly applied in the production of fine chemicals and pharmaceuticals where normally high productivities can be sacrificed as, in contrast to bulk chemicals production, the market of fine chemicals favours high regio- and/or enantioselectivity. These are both fundamental characteristics when designing a new drug.[9,10] According to Turner, the use of biocatalytic reactions will continue its growth giving the increasing number of available enzyme catalysts.[11] Turner used biocatalytic steps in many novel synthetic pathways for the production of interesting and complex molecules.[11,12]

1.2 Enzymatic and chemo-enzymatic cascade reactions

One of the advantages of enzymes is that they generally work under similar conditions as the result of millions of years of evolution under similar selective pressure. After all, in the cytosol of any living organism, thousands of enzymes work in combination with each other forming the whole metabolism of a given cell. Thus, the integration or combination of different enzymatic reactions requires limited tweaking of reaction conditions. This

enables the construction of synthetically interesting pathways for the production of chemicals of high molecular complexity.[13] Furthermore, the combination of different reactions in one step represents numerous advantages compared to the classic approach as it typically shortens the operation time, reduces the waste and improves the overall yields (Figure 2).[14] Indirectly, it enables better handling of harmful or unstable intermediates produced in certain chemical processes, as they are directly converted by the next catalyst of the cascade.[15]

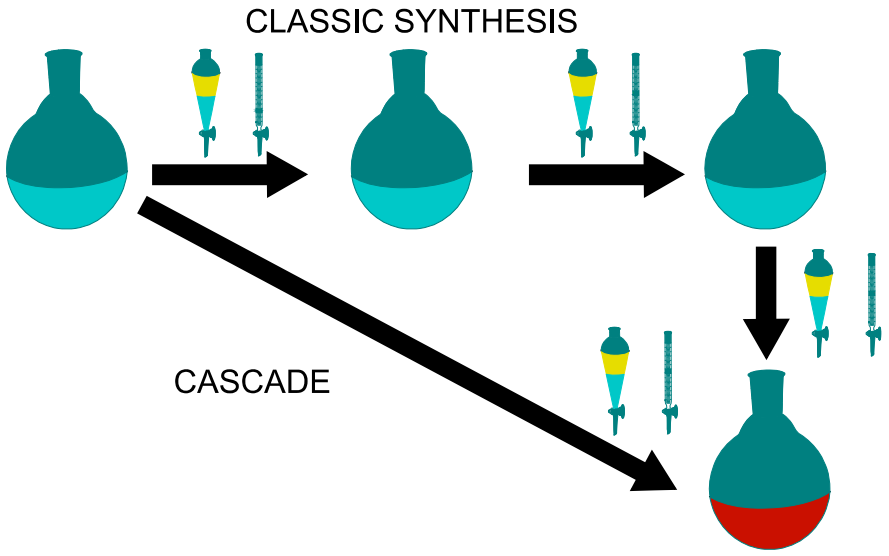


Figure 2. Schematic representation of the advantages of cascades over the classic syntheses. In the cascade approach, all steps are carried out without intermediate purification steps.

In the last years, the use of cascade reactions has become very popular and different classifications can be found. Four main designs can be defined depending on how the steps are running during the reaction course.[15] I) The most intuitive case is represented by the production of a single compound through multiple catalytic steps starting from a single substrate (Figure 3, A). II) The second type is widely applied and typically characterises enzymes that require a cofactor for their function. This cascade is defined as orthogonal as the main reaction is coupled with a secondary reaction to regenerate the cofactor (Figure 3, C). III) The third kind of cascade reaction is very similar to the second type, where two enzymes convert distinct substrates to produce different products. In this case, the reactions are also linked by opposite cofactor usage although both products are going to be isolated as they are of high value, whereas in the type II cascade, the co-substrate and co-product are discarded (Figure 3, B). IV) Finally, the fourth type, generally used for deracemization, can be considered a cyclic cascade where the enzyme only converts one of the enantiomers to an intermediate that is

converted back to the original mixture until only the non-reactive substrate enantiomer remains[15] (Figure 3, D).

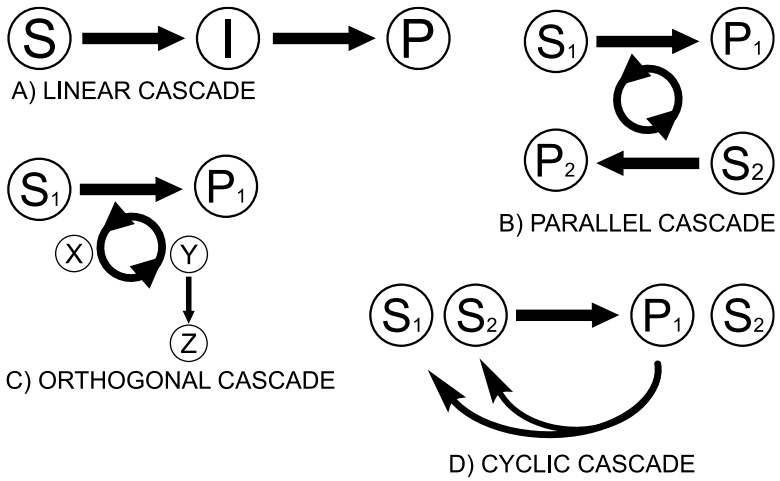
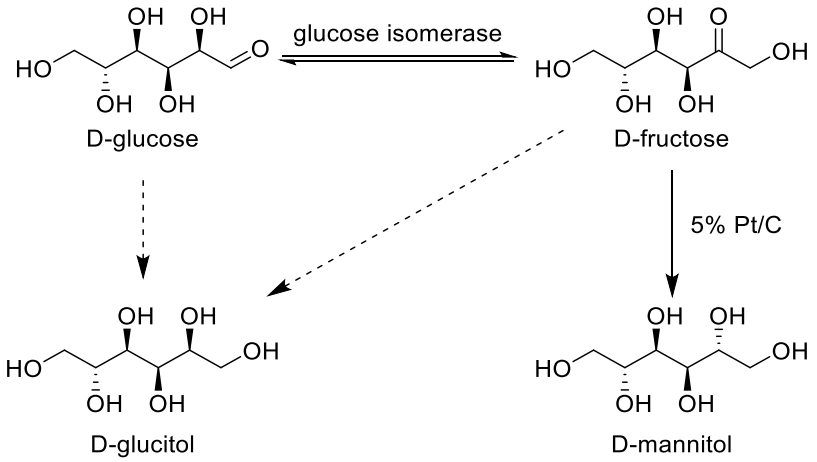


Figure 3. Four main types of cascade reactions (adapted from Ricca *et al*.[15]).

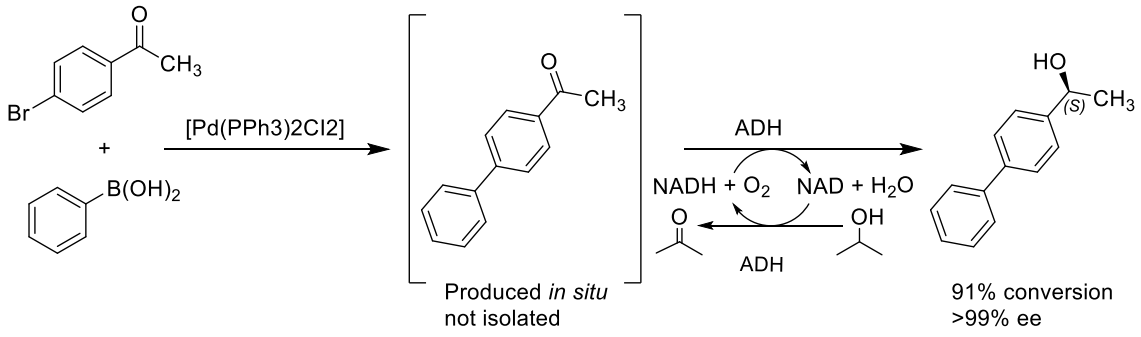
The different cascade designs can be combined to form new kinds and therefore the classification is not absolute. Moreover, cascades do not need to be strictly enzymatic or chemical. Instead, the synthetic chemist can pick from either enzymatic or chemical catalysts to assemble new synthetic routes. These two types of catalysts, in fact, were never exclusive but complementary to each other. The first example of a chemo-enzymatic cascade can be dated back to 1980 when a glucose isomerase was combined with platinum on carbon. The use of glucose isomerase was the key to achieve a fast interconversion of glucose to fructose while limiting the formation of the side product D-glucitol. As a result, the combination of these two catalysts could improve the D-mannitol yield from 25% to 46% (Scheme 1).[16]



Scheme 1. General scheme of the first chemo-enzymatic cascade where a glucose isomerase and platinum on carbon were combined to improve efficiency over a purely chemical process (Adapted from Makkee *et al.* [16]).

A decade later, cascade strategies gained more attention as they allowed overcoming the theoretical maximum yield of 50% in classic kinetic resolutions of racemates. This was achieved by combining a lipase with a palladium-catalysed racemisation resulting in a so-called dynamic kinetic resolution.[17,18] Lipase and palladium could perform the corresponding reactions simultaneously as no compatibility issues were observed.

In a more recent chemo-enzymatic cascade a palladium-catalysed Suzuki-cross coupling was combined with an alcohol dehydrogenase-catalysed reduction (Scheme 2). This time, the reactants of the first step were incompatible with the enzyme used in the second step.[19] The strategy used to solve observed compatibility issues was to run the steps sequentially, the so-called one pot two steps strategy. In the first step, a palladium complex is used to catalyse the Suzuki-cross coupling reaction and, once all the reagents are consumed, the (*S*)-selective alcohol dehydrogenase from *Rhodococcus sp.* is added to perform the dehydrogenation together with isopropanol for cofactor regeneration (Scheme 2).[19]



Scheme 2. Concept of the chemo-enzymatic cascade to combine Suzuki-cross coupling reaction with a biocatalytic reduction (adapted from Burda *et al.* [19]).

While enzymes typically work well together, the combination of chemical and enzymatic catalysts is not trivial. Often compatibility is achieved via extensive adjustment of reaction conditions as well as careful selection of reagents to suit all catalysts. One popular way to overcome incompatibility is compartmentalization to physically separate the catalysts and possibly the reaction conditions.[14] Despite the predictable difficulties, the design and set up of chemo-enzymatic cascades is becoming a popular approach to tackle the synthetic challenges that arise when a complex molecule needs to be synthesised.

1.3 Chiral azido and amino alcohols

The chemistry of compounds containing an azido group has caught the attention of organic chemists already in the previous century and they are still important building blocks for the synthesis of amino alcohols. Furthermore, they are key intermediates in the synthesis of carbohydrates and nucleosides.[20,21] Azido alcohols are generally synthesised by azidolysis of epoxides in aqueous alcohol systems. This type of reaction, however, requires high temperatures and long reaction times and no chemo- or stereoselectivity is observed.[21] Few chemical catalysts were designed to catalyse an enantioselective epoxide ring opening, however, their molecular complexity and price are major limitations.[22,23] In contrast, halohydrin dehalogenases (HHDHs)[24] and epoxide hydrolases[25] can perform regio- and enantioselective epoxide ring opening reactions and especially HHDHs can use sodium azide to produce azido alcohols in a selective way (Section 1.5).[26,27]

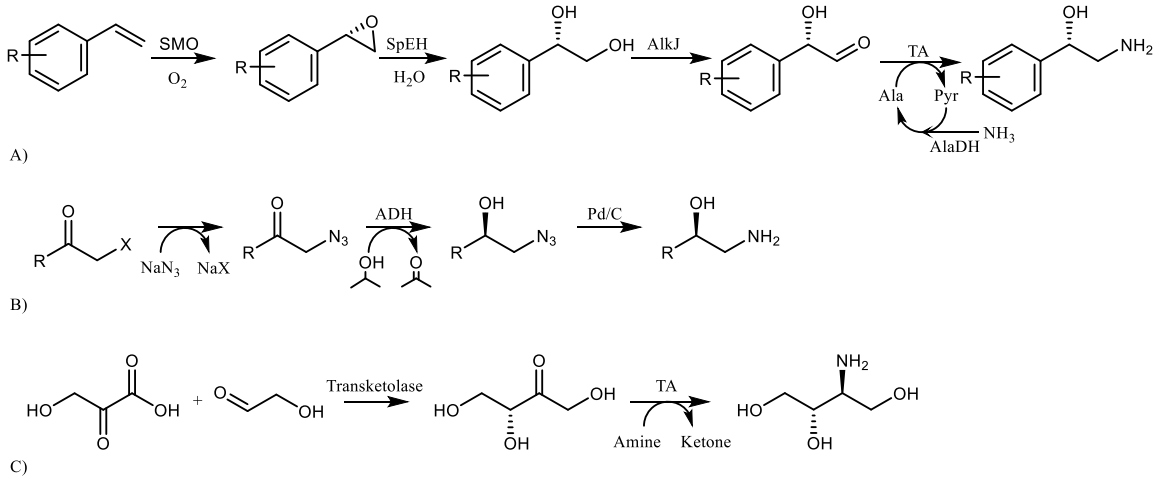
Amino alcohols are very important building blocks as well and they can be found in a number of natural compounds and drugs.[28] They are often used as chiral intermediates or chiral auxiliaries in organic synthesis. Furthermore, the simple manipulation of the hydroxyl- or amino-group can lead to a variety of other compounds.[29]

The synthesis of chiral amino alcohols starting from chiral amino acids is a common and simple synthetic route as only the reduction of the carboxylic group is required. Whereas a less common synthetic strategy is the intensive manipulation of carbohydrates that leads to amino alcohols.[30]

Most examples for the synthesis of chiral amino alcohol deal with terminal amino alcohols for which only one stereocenter needs to be controlled. In one example (Cascade A, Scheme 3), the cascade was composed by a styrene monooxygenase in the first step for the enantioselective epoxidation. The configuration of the hydroxyl group of phenylethanolamine and derivatives is controlled in the first reaction and the remaining steps are required to insert the terminal amino group. In particular, the styrene oxide is

hydrolysed by an epoxide hydrolase, consequently the terminal hydroxyl group is oxidised by an alcohol dehydrogenase. Finally, the amino group is inserted using a transaminase.[31] The approach is very interesting; however, reducing the steps could result in a more attractive process from an economic point of view as three steps out of four are used to insert the amino group.

Alternatively, phenylethanolamine could be synthesised starting from 2-chloro-1-phenylethanol (Cascade B, Scheme 3) reducing the number of steps to three compared to cascade A.[32] In this case, the halogen is removed via azidolysis and the stereocenter is controlled by reducing the keto-group in the second step using a selective ADH. Lastly, the azide is reduce to amino group and, as for cascade A, a terminal amino alcohol is produced. If a non-terminal halo ketone would be used, the configuration of the stereocenter where the amino group is attached would be predetermined by the chirality of the carbon-halide bond. Cascade C shows an example of the synthesis of a non-terminal amino alcohol starting from hydroxypyruvate and glycolaldehyde. Here both stereocenters are controlled during the reaction. In the first step, the transketolase controls the configuration of the carbon atom to which the hydroxyl group is attached and subsequently, a transaminase generates the second chiral centre.[33] Due to their natural function, transketolases can only be used for the synthesis of carbohydrate-like compounds, hence unfunctionalised compounds such as aliphatic alkenes would not be converted.[33,34]

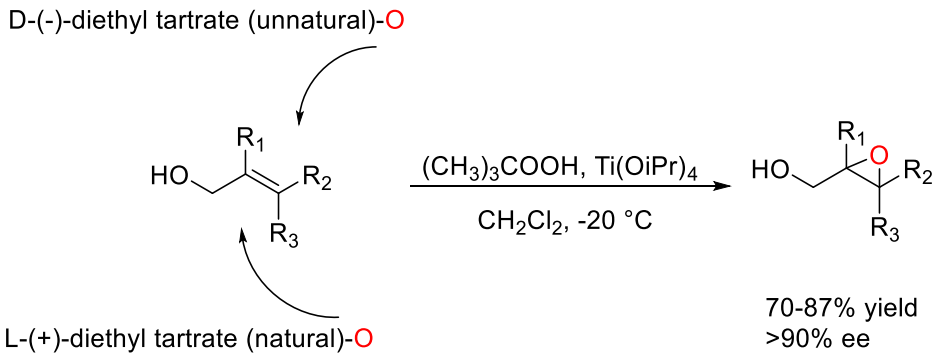


Scheme 3. Different cascade examples for the biocatalytic synthesis of chiral amino alcohols (adapted from Wu *et al.* (A),[31] Schrittwieser *et al.* (B),[32] and Ingram *et al.* (C)[33]).

1.4 Asymmetric epoxidations

1.4.1 Chemical catalysts

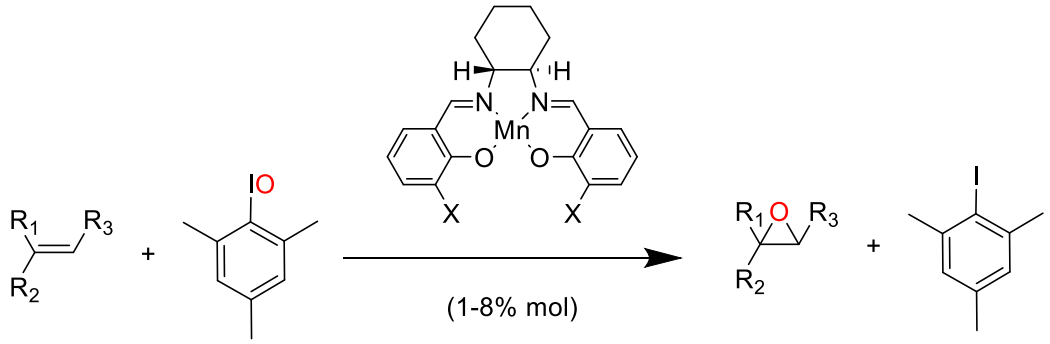
Several chemical catalysts have been developed to catalyse asymmetric epoxidations and the most notable example is certainly the Sharpless epoxidation. Sharpless, for the first time, could produce a chemical catalyst for the epoxidation of allylic alcohols that could reach up to 90% enantiomeric excess (Scheme 4).[35] His strategy was to use a chiral environment around the actual catalytic moiety, the transition metal (Mo, Ti or V). This discovery highlighted the importance of surrounding the catalytic metal with a chiral ligand to improve the enantioselectivity of the reaction. However, the first generation ligand was not optimized and had a low turnover number. Additionally, it was very specific for allylic alcohols while not performing well with homoallylic alcohols.[36,37] Several optimizations have been performed to improve the stability and enantioselectivity of the catalyst by changing the ligand, the transition metal and the oxygen donor. Nonetheless, the Sharpless epoxidation is very specific for allylic alcohols as the substrate's hydroxyl group is used as a pivot point.[38,39]



Scheme 4. General scheme of the first Sharpless epoxidation example (adapted from Katsuki and Sharpless[35]).

Salen complexes represent a different class of catalysts for the asymmetric synthesis of epoxides. In this case, Jacobsen and Katsuki paved the ground for what is now known as Jacobsen-Katsuki epoxidation.[40] The enantioselectivity in this reaction is controlled via a chiral environment as well but the epoxidation requires only simple oxidation agents such as PhIO, NaClO, H_2O_2 or Oxone® compared to titanium tetraisopropoxide required in the Sharpless epoxidation.[41] The synthesis of the chiral salen ligand is performed by condensation of diamines with salicylaldehyde derivatives allowing a substrate specific tuning of the substituents.[42] Originally, the salen complex was used to coordinate manganese (III) (Scheme 5), but in the later development chromium, cobalt, palladium, nickel and ruthenium were also applied.[38,43] The major limitation of this class of catalysts

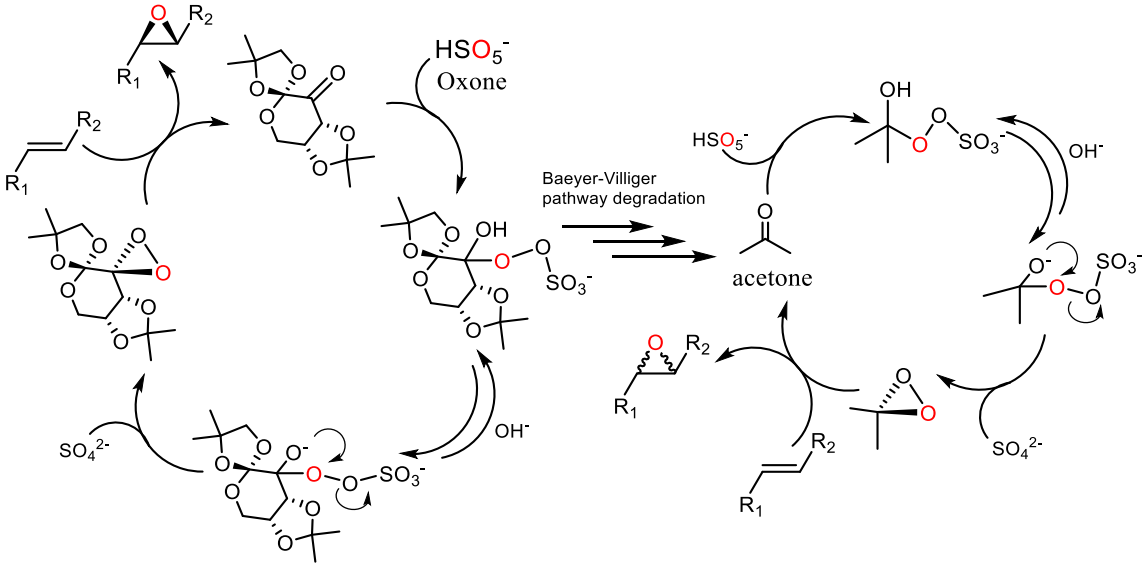
is the fact that they show high selectivity mostly towards *cis*-di-substituted olefins[44], however, examples for epoxidation of tri- and tetrasubstituted olefins (eg. 1,1-diphenylpropene) are also reported to show good performances, whereas selectivity towards *trans*-olefins was found to be low.[45,46]



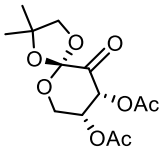
Scheme 5. General scheme of the first example for epoxidation of unfunctionalised olefins catalysed by (salen)manganese complexes (adapted from Jacobsen *et al.*[40]).

An additional method for asymmetric epoxidation is the use of dioxiranes. These moieties can be directly added to the reaction mixture, however, the most common method is the *in situ* synthesis using $KHSO_5$ (Oxone) and ketones. The first example was reported in 1984 where the epoxidation of (*E*)-β-methylstyrene and 1-methylcyclohexene was carried out with modest enantioselectivity (16-35% ee).[47] The most important example of the use of dioxiranes in asymmetric epoxidations, however, is the Shi epoxidation firstly reported in 1996.[48] In this case, the dioxirane is generated from a fructose-derived ketone which allows the reaction to proceed at mild conditions reaching extremely high enantioselectivity (Scheme 6). Moreover, the chiral centres in this catalyst are closer to the active ketone moiety ensuring a better stereoselectivity compared to other scaffolds holding keto groups. Additionally, the carbohydrate-derived scaffold is a rigid structure which is preferable to enhance enantioselectivity.[38] This fructose-derived ketone (Shi diketal epoxidation catalyst) shows optimal performances with *trans*-unfunctionalised olefins and poor selectivity with *cis*-olefins, in contrast to the Jacobsen catalysts ((salen)manganese complexes).[48,49] The major disadvantage of this catalyst is that it can undergo auto-Baeyer-Villiger oxidation producing acetone. The dioxirane can then be generated from the keto-group of acetone to carry out the same oxidation cycle but with no stereoselectivity (A, Scheme 6). This was partially solved by running the reaction at low temperatures and pH above 11 to prevent catalyst decomposition. Later, an acetate analogue of the original catalyst was synthesized. The acetate, with its higher electron withdrawing effect, could reduce the auto-oxidation making the catalyst more stable (B, Scheme 6).[50]

A) Shi epoxidation diketal catalyst reaction mechanism



B) Acetate analogue:



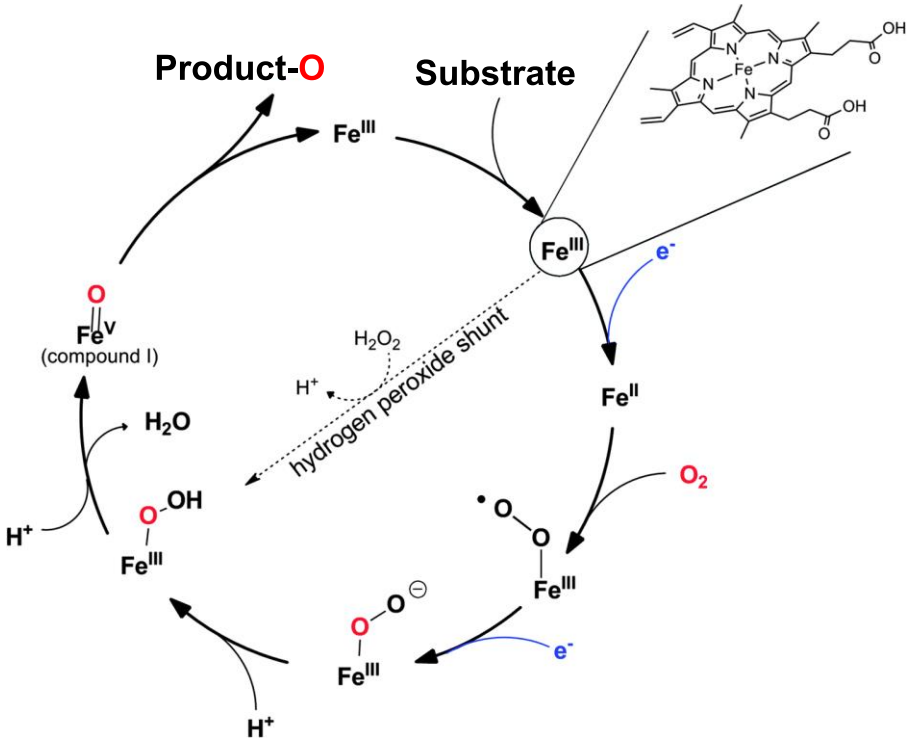
Scheme 6. A) General scheme of the epoxidation using the Shi epoxidation diketal catalyst and degradation pathway. (adapted from Shi *et al*[48]). B) Acetate analogue of the Shi epoxidation diketal catalyst.

1.4.2 Enzyme Catalysts for chiral epoxidations

Despite the available chemical catalysts for asymmetric epoxidation mentioned before, controlling stereo- and regio-selectivity in some alkenes is still challenging and in most cases, toxic metals and reagents need to be used.[51] Nature has evolved a few oxidative enzymes to perform asymmetric epoxidations on olefins using molecular oxygen. The most popular are monooxygenases, oxidases and peroxygenases.[52–54] Among these, cytochrome P450 monooxygenases (CYP450) carry a heme-iron prosthetic group for oxygen activation (Scheme 7).[55,56] Cytochrome P450s became very popular for their ability to hydroxylate non-activated carbon atoms with extremely high regio- and stereospecificity.[57] Being the main xenobiotic metabolizing enzymes in the human body, CYP450 research is often associated with studies on their interactions with drugs and other xenobiotics *in vitro*.[58,59] Lately, the advances of protein engineering and process engineering enabled also the application of these enzymes in the production of fine chemicals, fragrances and in bioremediation.[57] One of the oldest industrial examples is the regio- and stereoselective hydroxylation of Reichstein S at 11β-position to

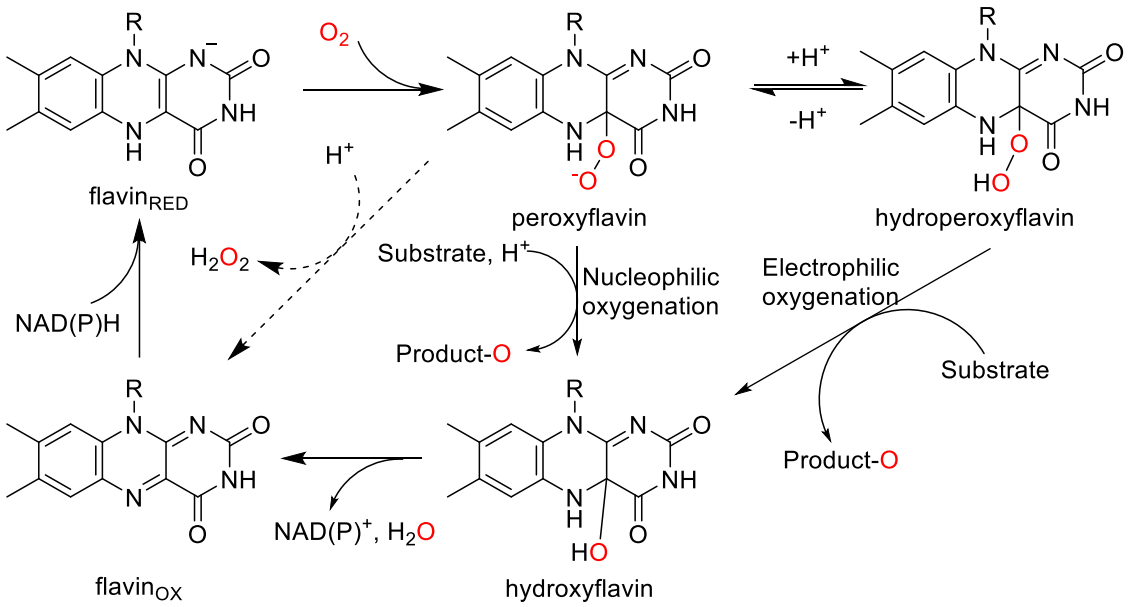
hydrocortisone using a CYP450 from *Curvularia* sp., which reached a production of approximatively 100 tons per year.[60]

Peroxygenases are similar to cytochromes P450s as they share the same prosthetic group and they can catalyse a similar range of reactions. On the other hand, they take advantage of the hydrogen peroxide shunt (Scheme 7) making them independent from the electron transfer systems that characterise cytochromes P450s.[54,61] Through the peroxide shunt pathway, the catalytic cycle of the Iron in the heme is shortened as compound I if formed using oxidants (e.g. hydrogen peroxide) instead of O_2 reducing the number of steps in the catalytic cycle. For approximatively 20 years, the chloroperoxidase from *Caldariomyces fumago* (CPO) was the only representative of this group of enzymes. Only recently, more peroxygenases applications have emerged due to their highly interesting chemistry.[62] The aromatic peroxygenase from *Agrocybe aegerita* (AaeAPO) is currently the more investigated and is at least twofold more active than CPO, furthermore it displays a broader substrate scope. Hence, it was classified as unspecific peroxygenases.[62] The major challenge in using peroxygenases in biocatalysis is the efficient hydrogen peroxide supply. On the one hand, hydrogen peroxide is essential for enzyme activity, on the other hand, high concentrations of this molecule cause irreversible degradation of the heme prosthetic group.[61,63] In the close future, it is possible that this class of enzymes could replace cytochrome P450s as they retain the same interesting chemistry but they use a simplified catalytic cycle.[64]



Scheme 7. Catalytic cycle of cytochromes P450 monooxygenases. The two electrons coming from NAD(P)H are transferred separately to the iron ion by the electron transport chain (adapted from Hollmann *at al.*[54]).

Flavin-dependent monooxygenases are another interesting group of enzymes that can perform epoxidations. These enzymes contain a flavin-cofactor to activate molecular oxygen and perform oxidations. In this reaction mechanism, the oxygen is activated before the substrate binds forming the reactive peroxyflavin.[65] Once it reaches the active site, the substrate is oxidised, typically with high regio- and stereoselectivity.[66] Next to epoxidations, they can catalyse hydroxylations, Baeyer-Villiger oxidations and sulfoxidations depending on the shape of the active site and exposed amino acid side chains.[65]

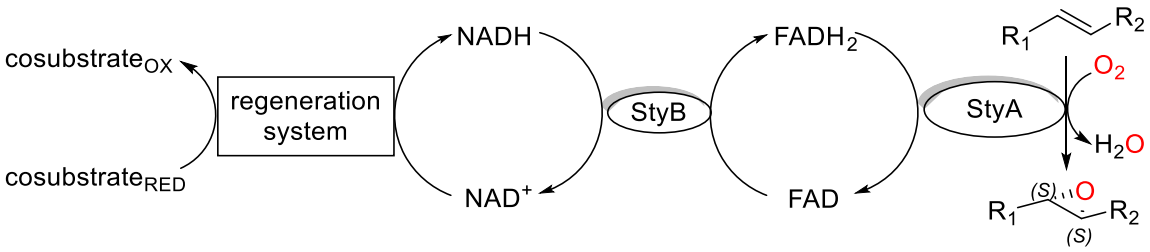


Scheme 8. Catalytic cycle of flavin-dependent monooxygenases. flavin$_{RED}$ and flavin$_{OX}$ represent the reduced or oxidised state of the flavin-cofactor, respectively (adapted from Fraaije *et al.*[56,65]).

Among flavin-dependent monooxygenases, Baeyer Villiger monooxygenases (BVMOs),[67] and styrene monooxygenases[68,69] are probably the most known together with flavin monooxygenases (FMOs). FMOs are a sub-class of flavin-dependent monooxygenases that carry a covalently bound cofactor.[70] BVMOs are probably the most familiar in organic chemistry as few multi-kg scale applications are already available (generally using whole cells).[71,72] The best known BVMO is cyclohexanone monooxygenase (CHMO) from *Acinetobacter calcoaceticus* which proved to have a wide substrate scope and high regio- and stereoselectivity.[71,73]

Styrene monooxygenases are getting increasing attention for their extremely high stereoselectivity in the epoxidation on unactivated C=C bonds. More specifically, the substrate scope is generally limited to vinyl aromatic compounds with few exceptions.[68,74] The styrene monooxygenase from *Pseudomonas* sp. VLB 120 is

probably the most characterised.[75–78] This enzyme could produce up to 388 g of styrene oxide in a biphasic system.[76] Interestingly, all known enzymes produce preferentially the (*S*)- or (*S,S*)-epoxide enantiomer despite several styrene monooxygenases from different sources have been characterized.[74,79–81] This, very likely, represents a temporary limitation and screening of more natural sources and/or engineering of existing enzymes should allow improvements of the substrate scope of these enzymes. Generally, styrene monooxygenases are composed of two subunits: the flavin-dependent catalytic subunit (StyA) and the flavin oxidoreductase (StyB) (Scheme 9), but there are examples of self-sufficient styrene monooxygenases with the two subunits fused together.[82] A recently characterised styrene monooxygenase from *Rhodococcus sp.* ST-10 was found to be able to convert short-chain aliphatic alkenes with good performance.[83] Toda *et al.* have shown that StyB from *Rhodococcus sp.* ST-10 is strictly dependent on NADH and no activity could be measured with NADPH, whereas it could accept both FAD and FMN. On the other hand, StyA from *Rhodococcus sp.* ST-10 exclusively accepted FAD for catalytic activity. Unfortunately, this enzyme displays substrate and product inhibition exhibiting the highest activity at only 2.5 mM styrene concentrations and when < 5 mM styrene oxide is present.[83] The inhibitory effect of substrate and product could be reduced by using a two-phase system with a second organic phase as found by Panke *et al.* using the styrene monooxygenase from *Pseudomonas* sp. VLB 120.[84] A second organic phase acts as substrate and product sink decreasing the toxicity for the biocatalyst.



Scheme 9. General reaction scheme for alkene epoxidation catalysed by styrene monooxygenases with optional cofactor regeneration system (adapted from Toda *et al.*[85]).

1.4.3 Cofactor regeneration strategies

Generally, enzymes that can activate molecular oxygen (e.g. monooxygenases) must rely on the availability of NAD(P)H to provide electrons needed for the catalytic mechanism; with the exception of peroxygenases that directly use hydrogen peroxide for substrate oxidation.

Adding NAD(P)H in stoichiometric amounts would result in an extremely expensive process, hence a few methods for *in situ* cofactor regeneration were set up. This allows using NAD(P)H in catalytical quantities.[53]

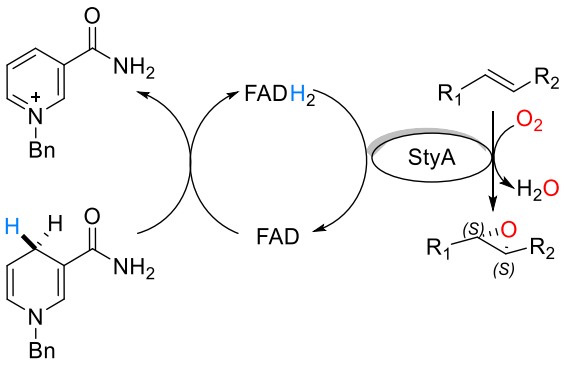
There are two main strategies to regenerate cofactors: the first one relies on the expression host's metabolism. This way, the cells provide NAD(P)H upon degradation of a cheap carbon such as glucose. A second strategy is to supply a cofactor regenerating enzyme that uses a sacrificial inexpensive cosubstrate that is oxidised to regenerate NAD(P)H (commonly used enzymes are listed in Table 1).

Using the host cell's metabolism is an easy method for *in situ* cofactor regeneration, as no coexpression of additional enzymes is required. However in some cases, the substrate or product can be toxic for the cells resulting in inefficient NAD(P)H supply. Here, using a cofactor regeneration enzyme would make the reaction independent from cellular metabolism. Furthermore, by not relying on cell's metabolism, the enzyme catalysts can be applied as cell-free extract improving the reaction rate as some substrates cannot diffuse easily through the cellular membrane.

Table 1. Common cofactor regeneration enzymes that are used in biocatalysis in combination with monooxygenases (adapted from Holtmann *et al*.[54]).

Enzyme	Cosubstrate	Coproduct	Cofactor	Example
FDH	Formate	CO_2	NADH	Churakova *et al.*[86]
ADH	Isopropanol	Acetone	NAD(P)H	Schulz *et al.*[87]
PDH	Phosphite	Phosphate	NAD(P)H	Watanabe *et al.*[88]
G6PDH	Glucose-6-phosphate	Glucono-6-phosphate lactone	NAD(P)H	Rioz-Martínez *et al.*[89]

An additional strategy developed to tackle the high costs of the cofactors NADH and NADPH is to chemically synthesise and use cheaper analogues instead. In these cofactor mimics, the adenosine dinucleotide is exchanged with a shorter synthetic substituent. They are known since 1936,[90] however, only recently they received more attention for applications in biocatalytic reactions. Their redox potential can be tuned to spontaneously reduce FAD[91] and being cheap, they can be used in stoichiometric amounts.[92] In one recent example by Paul *et al*., 1-benzyl-1,4-dihydronicotinamide (BNAH) was used as NAD(P)H mimic in the epoxidation of styrene catalysed by the styrene monooxygenase from *Rhodococcus opacus* 1CP. In this specific case, two enzymes (regeneration system and StyB, Scheme 9) could be omitted due to the use of BNAH to reduce the flavin needed by the catalytic subunit (StyA, Scheme 10).[91,92] Despite the clear advantages, real applications of cofactor mimics in biocatalysis are still scarce.

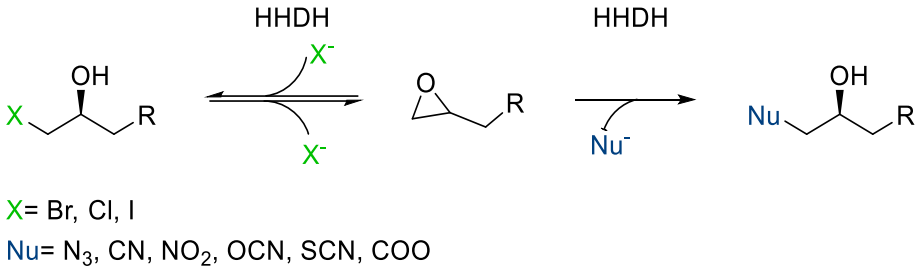


Scheme 10. *In situ* regeneration of $FADH_2$ using the cofactor mimic BNAH. Used in stoichiometric amounts, BNAH substitutes for StyB and cofactor regeneration system in the epoxidation catalysed by the styrene monooxygenase (see Scheme 9, adapted from Paul *et al*.[91]).

1.5 Epoxide ring opening

Epoxide ring opening is a common reaction that has been used in organic synthesis for many years. However, the chemical reactivity of the epoxide ring is rather low and depends on the surrounding if more or less tension on the ring is present.[93,94] The chemical ring opening usually does not proceed with regio- or stereoselectivity and, if any, the selectivity is substrate-controlled. Satisfying regioselectivity was only reported for terminal or aromatic epoxides, whereas the ring opening of linear epoxides carrying similar substituents on both sides of the ring proceeds with poor selectivity.[95] Currently, most of the chemical catalysts that would allow regio- or enantioselective ring opening require a very challenging synthesis and showed high performances only with terminal or meso-epoxides.[22,96]

On the other hand, a group of enzymes called halohydrin dehalogenases (HHDH, EC 4.5.1.-) can catalyse the epoxide ring opening under mild conditions and with chemo-, regio- and enantioselectivity depending on the substrate. The natural function of HHDHs is actually the reversible dehalogenation of vicinal haloalcohols forming the corresponding epoxide.[26,97–99] The dehalogenation, a reversible reaction, can be exploited by providing different nucleophiles such as cyanide, azide, nitrite, cyanate, thiocyanate or formate to catalyse the irreversible epoxide ring opening (Scheme 11).[24,100,101]



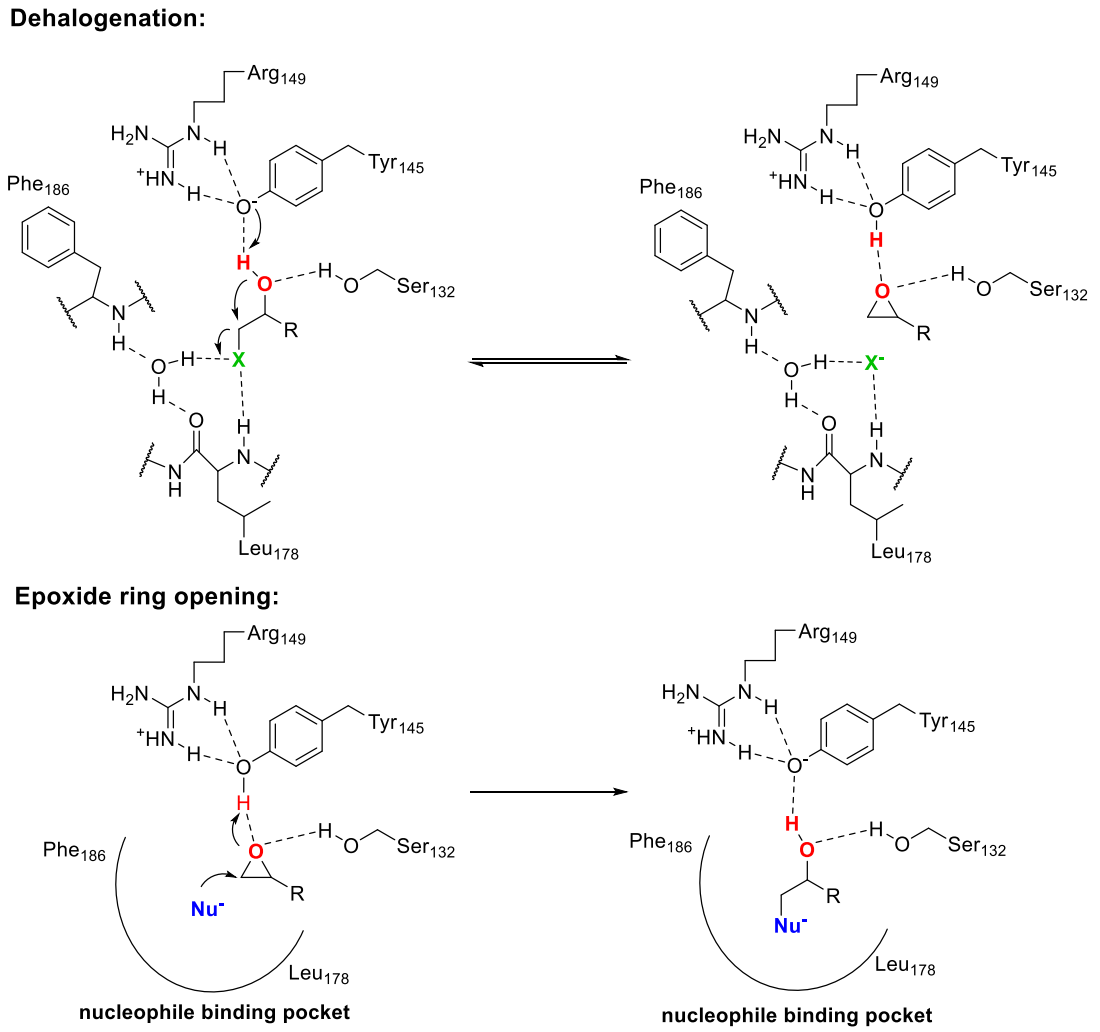
Scheme 11. Reaction scheme of reversible dehalogenation and irreversible epoxide ring opening catalysed by halohydrin dehalogenases.

These enzymes are both structurally and mechanistically related to the short-chain dehydrogenase/reductase (SDR) superfamily. HHDHs assume a tertiary structure that resembles a Rossman fold and they show a quaternary structure of four subunits (homotetramer), these structural features are also typical of SDR-like enzymes. Furthermore, the Ser-Tyr-Arg catalytic triad in HHDHs has a similar function as for the Ser-Tyr-Lys catalytic triad in most SDR enzymes. The main difference is that HHDHs evolved a halide/nucleophile binding site instead of the nicotinamide cofactor binding site found in SDR enzymes.

Recently, the number of available HHDH sequences was extended through database mining to 65.[102,103] These new HHDH sequences were identified in public databases, using a novel approach, by filtering out all SDR-like sequences that did not present the HHDHs catalytic triad motif (S-X12-Y-X3-R) and the nucleophile binding pocket motif (T-X4-F/Y-X-G).[102] Noteworthy, the phylogenetic analysis of the newly discovered sequences revealed that most of them could not be classified in the existing phylogenetic groups (A, B and C) and four new subtypes were added (D to G).[102] A subset of 17 from each subtype were then isolated and extensively characterized to prove that they were true halohydrin dehalogenases.[104,105]

The active site assumes a tunnel-like shape on most HHDHs and the shape is enzyme dependent and residues composing the tunnel are known to influence activity and selectivity, as demonstrated in HheC.[106–109] HheG, in contrast, shows an open cleft that possibly confers it the broad substrate scope.[105] In the dehalogenation reaction, the catalytic triad function is to abstract the proton of the hydroxyl group present in the substrate. In SDR enzymes, the negative charge left on the hydroxyl oxygen is transferred to the $NAD(P)^+$ cofactor; in HHDHs the oxyanion, stabilised by the serine, performs a nucleophilic attack on the carbon where the halogen is present. As a consequence, the halogen is displaced leading to the epoxide formation (Scheme 12).[110] Conversely, in the epoxide ring opening reaction, the halide (or one of the accepted nucleophiles) performs a nucleophilic attack on one of the carbons of the

epoxide ring leading to cleavage of the carbon-oxygen bond. The tyrosine donates a proton to the oxyanion leading to the formation of the halo-alcohol or other β-substituted alcohols (depending on the nucleophile).



Scheme 12. General dehalogenation and epoxide ring opening reaction catalysed by the HHDHs. (Residue numbering according to HheC, adapted from Schallmey *et al.*[103])

Cyanide is possibly the most interesting nucleophile accepted by HHDHs as it allows the formation of a novel C-C bond, which is generally challenging in classic organic chemistry. The cyano-group can further be manipulated to access amino, amide, or carboxyl groups.[111,112] Azidolysis also represents an important reaction catalysed by halohydrin dehalogenases as it yields azido- and amino-alcohols (after a simple reduction of the azido group) that are important building blocks as described in section 1.3. In a process developed by Codexis Inc. (Redwood City, CA, USA)[113] a highly engineered HheC mutant was applied in the synthesis of the statin precursor. In this

process, a HheC mutant performs in two steps the dehalogenation of ethyl (*S*)-4-chloro-3-hydroxybutyrate to the intermedia epoxide and consequently, the epoxide ring opening using cyanide as nucleophile producing (*R*)-4-cyano-3-hydroxybutyrate.[114,115] Atorvastatin, also known as Lipitor, is a cholesterol-lowering drug which was the top-selling prescription medication until 2010 with more than $7 billion worth in total revenue.[116]

1.6 Hydrogenations

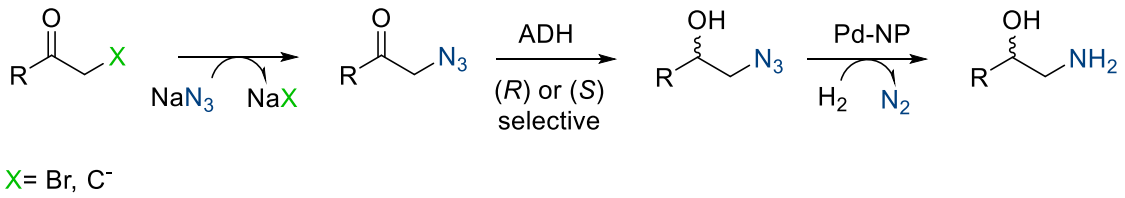
The ability of certain transition metals to react with molecular hydrogen has been known for roughly a century. The presence of these catalysts proved to be essential to promote hydrogenation of certain substrates at high temperature.[117] The reactivity of the metals can be tuned by different additives or supports and in the years, several catalysts were produced to fit the kind of compounds to be hydrogenated.[118] These metals can be used as salts, that can dissolve in solution (homogeneous catalysts). However, heterogeneous catalysts offer more advantages especially when it comes to applicability. They are called 'heterogeneous' as the metal ions are immobilised on organic or inorganic (or a combination of both) supports and they do not dissolve in solution. The main advantage here is the fact that the supports bring the size of the catalyst to a macromolecular level enabling easy separation by filtration during work up.[118]

The first industrial process, established in 1897, was based on a Nickel catalyst and since then a number of metals (Table 2) were applied on supports, as complex oxide systems or as Raney metal catalysts.[119]

Table 2. Common hydrogenation reactions and the respective metal catalysts used (adapted from Navalikhina *et al.*[119]).

Reaction	Catalyst
Hydrogenation of alkenes	supported Ni, Cu, Pt, Pd, Ru, and Rh
Hydrogenation of alkynes	supported Ni, Pd, Ni-Pd, Co-Pd, and Cu-Pd
Hydrogenation of aromatic hydrocarbons	supported Ni, Pd, Rh, Pt, Co-Mo, Raney Ni
Hydrorefining of motor fuels	Ni, Ni-Mo, and Co-Mo on Al_2O_3 and diatomaceous earth
Hydrogenation of carbonyl compounds	skeletal and supported Ni and Co, compounds supported Pd, Rh, and Pt, and Cu-Cr-O, Cu on diatomaceous earth
Hydrotreating (Hydrodesulfurisation)	Ni-Mo, Ni-W, Co-Mo, Co-W, and Ru on Al_2O_3 and carbon, Mo and W carbides
Hydrogenation of nitriles, azides	supported Ni, Pd, and Pt
Hydrogenation of (methanation)	supported Ni, Cu, Co, Fe, Pt, and Rh, Pt mesh, Mo and W carbides

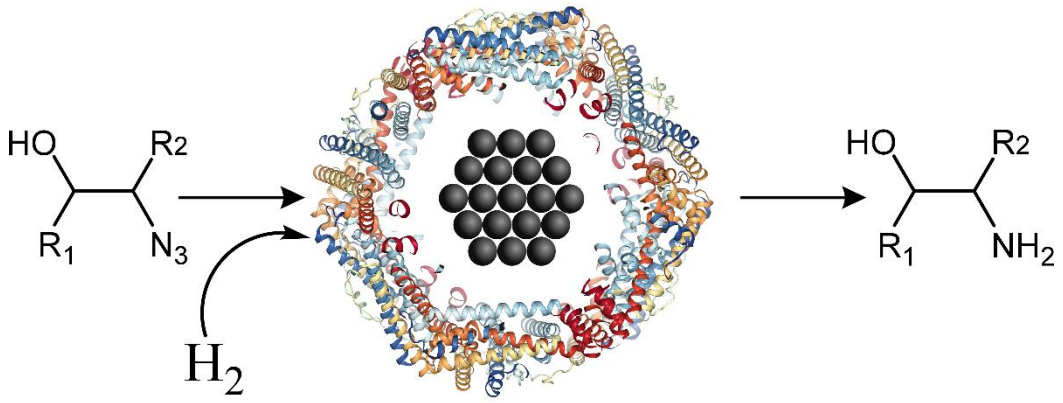
Nowadays, heterogeneous palladium is one of the most used catalysts to perform hydrogenation of most reducible functionalities due to its high activity and easy handling.[120] Heterogeneous palladium on carbon (Pd/C) is possibly the most used hydrogenation catalyst and due to its high reactivity, shows low chemoselectivity. The reactivity and consequently chemoselectivity can be adjusted in two ways: by catalysts poisons (*e.g.* sulphur) or additives (*e.g.* lead acetate) that can tune the reactivity. The second way to improve chemoselectivity is by selecting the characteristic of the support (*e.g.* surface area or pore size).[120] Typically, palladium catalysed hydrogenations are performed in ethanol or methanol in an autoclave reactor under pressurised H_2. Pd/C, thanks to its high reactivity, can be used in aqueous buffer as well with little to no activity loss. Recently, Schrittwieser *et al.* used lignin stabilised palladium nanoparticles to efficiently reduce 2-azido-1-aryl alcohols to 2-amino-1-aryl alcohols in phosphate buffer pH 9 under pressurised H_2 (Scheme 13).[32] The reduction of azide to amine occurred in four hours reaching full conversion in this case the chirality of the molecules was determined by the alcohol dehydrogenase (ADH) catalysed reduction as the terminal carbon is not chiral (Scheme 13). In any case, the reduction of azides to amine does not involve chiral centres, hence the chirality should be already present in the substrate (*e.g.* synthesised via azidolysis of an enantiopure epoxide[121]).



Scheme 13. Chemo-enzymatic cascade for the synthesis of enantiopure 1,2-amino alcohols using palladium catalysed azide hydrogenation (adapted from Schrittwieser *et al.*[32]).

As mentioned previously, Pd/C is one of the most used hydrogenation catalysts, however when combined with enzymes few challenges can arise. It is known that carbon nanoparticles and nanotubes can interact with the hydrophobic patches of proteins causing destabilization of the protein's tertiary structure. This results in the formation of a "corona" of unfolded peptides around the nanoparticles;[122] and, in the case of a Pd/C catalysed azide hydrogenation, the "corona" would act as a barrier between the substrate and the nanoparticles. Changing the support to one more "protein-friendly" is the more straightforward solution (*e.g.* lignin stabilised palladium nanoparticles[32]). An interesting strategy is to exploit the ability of ferritin to encapsulate the desired metal. Naturally, ferritin is used by organisms to safely store iron and use it when the intracellular iron

concentration is too low.[123] Thus, incubating ferritin in its apo-form with clusters of different metals including Pd would lead to the encapsulation of the metals inside the ferritin cage. Furthermore, Pd, for example, can be reduced to form nanoparticles of extremely uniform size inside the ferritin cage that act as a micro-reactor. This complex can then be used to perform hydrogenations without interfering with other protein or enzymes present in the reaction solution, as the nanoparticles are isolated by the ferritin cage; but at the same time, the substrate can freely diffuse inside the cage and be reduced[124] (Scheme 14). Palladium in ferritin has been successfully applied in analytical scale reactions, however, the long process to synthesise the Pd-ferritin nanoparticles has not yet found its way to industrial applications.[125]



Scheme 14. Example reaction scheme using azido alcohols as substrate and palladium nanoparticles encapsulated in a ferritin cage. The ferritin cage would prevent adsorption of external proteins onto the metal nanoparticles whereas substrate and hydrogen can freely diffuse inside the Pd-ferritin complex and react.

1.7 Protein engineering

Wild-type enzymes normally work at submillimolar concentrations in their natural functions. For biocatalytical application, however, significantly higher substrate concentrations are required hence, often the natural activity of a certain enzyme is not fitting for direct practical application in biocatalytical and industrial processes. Most of the times, it is necessary to improve the enzyme's activity and/or stability to fit the desired reaction conditions of a biocatalytic process.[5,126] Enzymes' function and structure are the results of their sequence, hence manipulating its sequence leads to variants with different characteristics. Since the advent of molecular technologies to efficiently manipulate DNA, enzymes have been evolved to improve certain characteristics,[7] to develop unnatural reactions[127,128] or to match industrial requirements.[129]

Mimicking natural evolution, a given enzyme can be mutated and selected for the desired characteristic iteratively until satisfaction. There are two main ways for "*in vitro*" enzyme evolution: directed evolution and rational design.

Directed evolution

The definition of directed evolution was used for the first time around 1972 and in the first experiments, whole cells were subjected to artificial environments that would cause "adaptive" changes in a given enzyme (Figure 4, A).[130] Francis and Hansche evolved and acid phosphatase from *Saccharomyces cerevisiae* by constructing a selection environment in a chemostat to reduce the enzyme's activity of about 70% (pH 6). Mutants with improved acid phosphatase activity at pH 6 would be positively enriched in the chemostat as β-glycerophosphate, metabolised by the enzyme, was provided as sole carbon source.[130] After 1000 generations of asexual reproduction, adaptive mutants were analysed to determine the new genetic and biochemical features. In the first part of the 1990s, Frances Arnold and Pim Stemmer improved this technique by using advanced molecular biology methods coupled with efficient screening methods to improve enzymes.[5] In directed evolution, only the enzyme's DNA sequence is required as the mutations are randomly introduced throughout its sequence to generate large mutant libraries. The most popular way to do that is the use of error-prone PCR (epPCR).[131] Mutagenesis and screening or selection can then be iterated until the enzyme displays improved characteristics or acquired the wanted features. The major drawback is the need for a screening method using a fast and reliable assay to efficiently identify improved variants before starting with the next round of mutagenesis (Figure 4, B).[132]

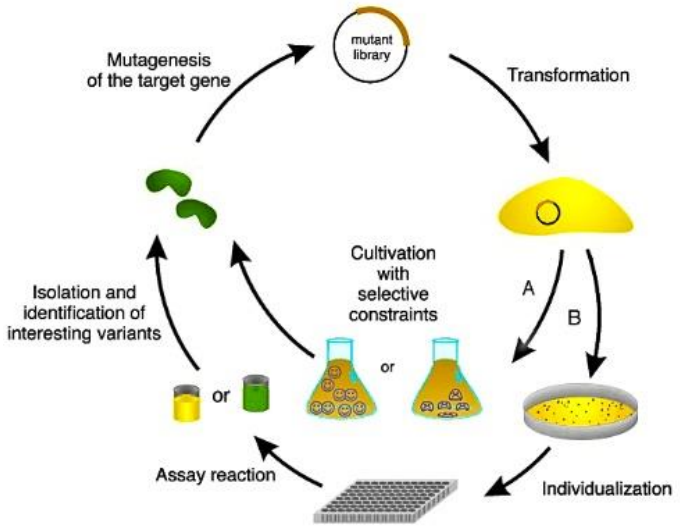


Figure 4. General representation of directed evolution using *in vivo* selection (A) or high-throughput screening (B) (adapted from Brakmann[133]).

Rational design

Unlike directed evolution, rational design strongly relies on the availability of structural information on the enzyme and knowledge about the mechanism/function. The exponentially increasing number of protein structures generated by X-ray crystallography, NMR, and cryo-electron microscopy made this approach more and more accessible.[134] Using available molecular modelling tools (*e.g.* PyMol or Yasara), it has been possible to "rationally" select residues that may be responsible for substrate specificity or enantioselectivity. This is generally done by replacing the single residues with amino acids present in homologous sequences.[135] Hence, the number of mutants to analysed is significantly reduced and there is no need for an efficient screening method.

Site-saturation mutagenesis (SSM), is a hybrid approach between random mutagenesis and rational design as selected positions are randomly replaced with all amino acids. Mutations are introduced on DNA level by replacing the codon coding for the specific amino acid with all possible codons (64 codons, 20 amino acids). SSM of a single site using an NNN-codon requires the screening of 192 random variants to achieve a 95% coverage considering an oversampling factor of three.[136] This number increases to 12'288 if two positions are mutated at the same time. Hence, smart libraries have been developed to decrease the number of variants to screen but keeping the 95% coverage. Due to the redundancy of the genetic code, the number of codons introduced during mutagenesis can be reduced from 64 to 32 (NNK-strategy with K coding for G or T) and still covering all 20 amino acids. This way, it is possible to reduce the number of variants to screen to 32 or 3'072 for 1 or 2 mutated positions, respectively.[132] In a later approach, the number of introduced codons was further reduced to 12 (NDT codon) by eliminating the chemical redundancy of amino acid residues, but keeping representatives of aromatic, aliphatic, non-polar, polar, negatively and positively charged amino acid residues.[136] Statistically speaking, the NDT-codon strategy is by far more efficient than the NNK-strategy (based on the screening of 5000 clones), and mutagenesis on the epoxide hydrolase from *Aspergillus niger* proved its validity by producing a library with 511 hits whereas the NNK-library only contained 38.[132,136] Furthermore, rational design allows a more in-depth characterization of the mutated position by testing different amino acids compared to random mutagenesis where generally a lower number of mutants can be explored at each position.[134]

Computational approaches

As mentioned above, the advent of DNA sequencing and manipulation has brought exceptional benefits to a number of fields like molecular and synthetic biology and biocatalysis. In parallel with molecular techniques, computational approaches

(bioinformatics) raised in popularity with the increasing availability of sequence data. In the past years, the sophistication of the algorithms has exponentially increased the ability to extract more and more useful information from this enormous amount of sequence data.[137] Moreover, a number of different algorithms have been developed to help to predict protein function based on the sequence and/or three-dimensional structure of a protein,[138,139] to predict mutations for improved stability[140] or enhanced substrate binding, or to completely change the substrate specificity of an enzyme.[141,142] The improved prediction efficiency significantly contributed to the growing success of these bioinformatics tools together with a user-friendly graphical interface. One of the most popular software packages for protein structure visualization, YASARA Structure,[143,144] has several built-in functions where a simple protein-ligand docking is only a few clicks away. Pymol, another famous software for protein structure visualization, has several plugins available that allow using bioinformatics methods without coding experience.[145] As mentioned previously, these bioinformatic tools can help to visualise the enzymes and select specific positions to mutate reducing the screening effort. Recently, a novel algorithm was implemented in the Rosetta software suite[146] here a new strategy was refined to predict mutations improving the interaction between proteins and small molecules. This new method, called “coupled moves”, takes into consideration the subtle rearrangements that can take place at the backbone level after a mutation.[147] Backbone rearrangements, side-chain conformational changes, amino acid mutations or ligand movements occur at the same time and the new calculated energy determines the probability of the mutation or conformational change (Figure 5). Once started, the algorithm changes residues selecting the amino acids that improve the binding energy with the ligand in a sort of *Darwinian way* mixed with a random acceptance-rejection rate, typical of the Metropolis-Monte Carlo algorithm.[148] This procedure is repeated for N times (determined by the user) and a list of accepted sequences is the output. This new approach could predict mutations improving protein-ligand interaction with a prediction efficiency up to 68% compared to the previous method with fixed backbone.[149] The method is very powerful although the enzyme’s reaction mechanism or a ligand-bound structure must be known as the new non-natural ligand is manually docked into the catalytic tunnel of the enzyme. The successful run must then be experimentally verified by generating the top scoring mutants and by checking their activity. This approach can also be combined with classic site-directed mutagenesis where the coupled moves method is used to identify target positions for SSM. If in the coupled moves output, a specific position appears to have high variability, it is very likely that mutations in that position will result in better protein-ligand interaction as other amino acids were accepted

more frequently than the wild type. Hence, a site where the wild type amino acid showed low frequency would likely be a potential target for SSM.

An additional and important advantage is the possibility to include epistatic mutations. Epistasis is the phenomenon where the effect of a given mutation becomes relevant only after other changes have taken place and often they appear to be neutral or deleterious for the protein to evolve.[150,151] Epistatic effects are normally neglected in classical protein engineering strategies as only positive variants are selected in screening methods. With the coupled moves strategy, some random mutations may be accepted and be beneficial at a later stage as multiple residues are mutated at the same time.

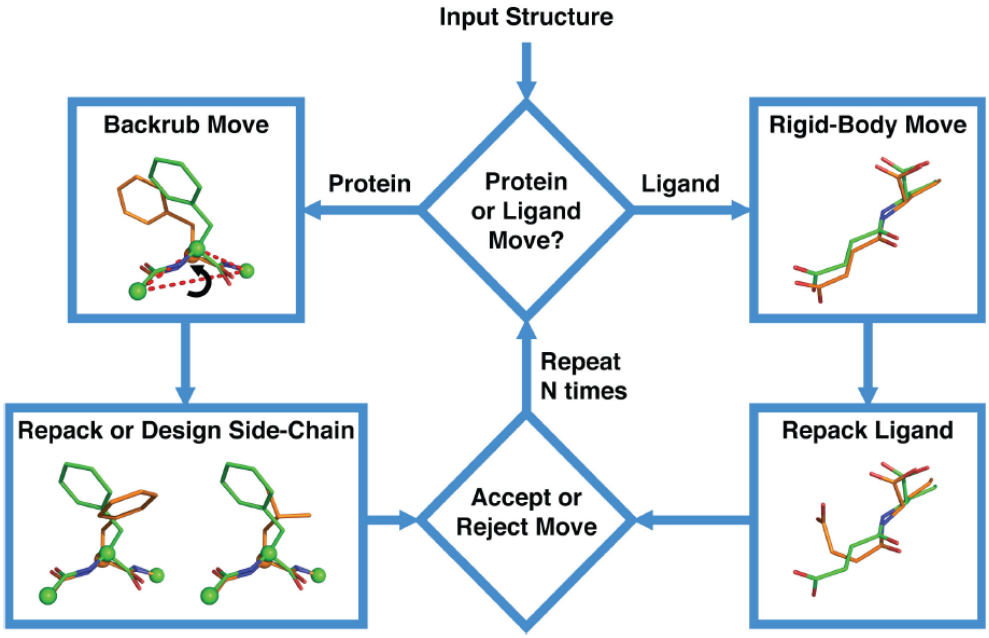
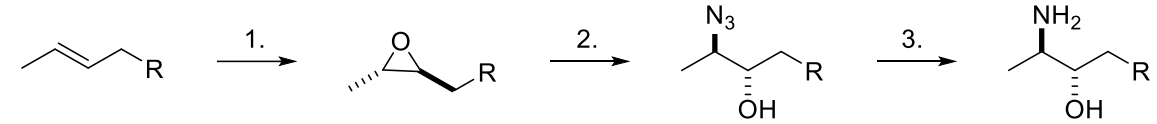


Figure 5. Flow-chart representation of the strategy used in the coupled moves algorithm implemented in the Rosetta software suite (taken from Ollikainen *et al.*[149]).

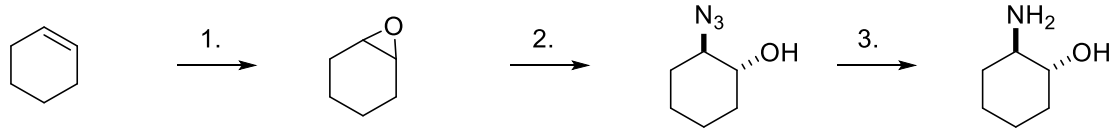
1.8 Aim of the project

The main objective of this project was to develop a new chemo-enzymatic route for the synthesis of enantiopure vicinal amino alcohols starting from non-terminal alkenes. For this, two main cascades were envisaged for the synthesis of (2*S*,3*R*)- and (2*R*,3*S*)-2-amino-3-hydroxyhexane as well as (2*S*,3*R*)- and (2*R*,3*S*)-2-amino-3-hydroxyheptane starting from achiral *trans*-2-hexene and t*rans*-2-heptene, respectively. Additionally, a third cascade was aimed for converting cyclohexene to the corresponding chiral 2-aminocyclohexan-1-ol. All cascades should combine the following three reaction steps: 1) epoxidation of the alkene, 2) epoxide ring opening using azide as nucleophile and 3) hydrogenation of the resulting azido alcohol (Scheme 15).



R = C_2H_5, C_3H_7

Cascades starting from acyclic alkenes



Cascade starting from cyclic alkenes

Scheme 15. General reaction scheme for the cascades starting either from non-terminal *trans*-alkenes for the synthesis of enantiopure acyclic amino alcohols or starting from cyclohexene for the synthesis of the corresponding cyclic amino alcohol.

Starting from a non-terminal alkene, the final amino alcohol will carry two stereocenters. To control the configuration of both stereocenters in the final product, the aim was to combine an enantioselective epoxidation (step 1) of the acyclic alkenes with a regioselective epoxide ring opening (step 2) to afford single enantiomers of the corresponding azido alcohols. In contrast, starting from cyclohexene will only require an enantioselective ring opening of the corresponding cyclohexene oxide as the epoxide is not yet chiral (Scheme 15). The final hydrogenation step should then proceed with preservation of the enantiopurity.

Hence, the first project objective was to identify suitable catalysts with desired selectivity for each reaction step by testing several available chemical and enzymatic catalysts. The subsequent project goal was to optimize each reaction step individually and to study possible compatibility issues in order to identify optimal conditions for the combination of the individual reaction steps in cascades. Finally, upscaling of the established cascades should be performed to isolate and characterize the formed products.

Additionally, for the cascade starting from cyclohexene, engineering of HheG was aimed for to improve its enantioselectivity using a new *in silico* method. This was necessary since HheG was the only HHDH able to convert cyclohexene oxide but displayed only moderate enantioselectivity in the ring opening of this epoxide using azide as nucleophile.

2. Materials and Methods

2.1 Materials

2.1.1 Chemical standards and substrates

All chemicals were purchased from either Sigma Aldrich Chemie (Steinheim, Germany), AppliChem (Darmstadt, Germany), abcr GmbH (Karlsruhe, Germany), Acros Organics (Geel, Belgium) and Carl Roth (Karlsruhe, Germany) unless otherwise stated, taking care of ordering the highest quality available.

Trans-2-hexene (**1**) and cyclohexene oxide were purchased from Acros Organics (Geel, Belgium), *trans*-2-heptene (**2**) was purchased from abcr GmbH (Karlsruhe, Germany). *Trans*-2-hexen-1-ol (**3**) was purchased from J&K Scientific bvba (Lommel, Belgium). *Trans*-3-heptene (**5**) and *trans*-β-methylstyrene (**5**) were purchased from TCI Deutschland GmbH (Eschborn, Germany). (*S,S)*-1-phenylpropylene oxide (**25**), (*R,R)*-1-phenylpropylene oxide (**25**), Shi epoxidation diketal catalyst, (*R,R*)- and (*S,S*)-Jacobsen's catalyst, formate dehydrogenase from *Candida boidinii,* apoferritin from horse spleen, heterogeneous Palladium catalysts kit I (2017), Pd/C, loading 10 wt%, 2-chlorocyclohexanol, (1*R,*2*R*)- and (1*S*,2*S*)-2-aminocyclohexanol, hydrazine hydrate solution 78-82%, triethylsilane (TES) were purchased from Sigma Aldrich Chemie (Steinheim, Germany).

Chemical synthesis of the substrate epoxides

The respective racemic epoxides *trans*-2,3-epoxyhexane (**21**) and *trans*-2,3-epoxyheptane (**22**) were synthesised using *meta*-chloroperbenzoic acid (*m*-CPBA) through a slightly modified protocol from Sharma *et al.*[152] Epoxidations were carried out in CH_2Cl_2 (40 mL/g of alkene). *m*-CPBA (1.5 eq.) was added in small portions over 15 min at RT. After 1 h at RT, the mixture was stirred on ice for 5 min and the white slurry was filtered. 5% w/v aq. Na_2SO_3 (40 mL/g alkene) was added and stirred at RT for 15 min, then the phases were separated. The aqueous layer was extracted two times with CH_2Cl_2. The combined organic layers were washed with sat. aq. $NaHCO_3$ (1×) and brine (1×), dried over $MgSO_4$ and filtered before solvent removal by evaporation. The epoxides were purified by column chromatography using a mixture of cyclohexane/ethyl acetate, 95:5.

The racemic epoxides *trans*-4,5-epoxyhexanol (**23**), *trans*-3,4-epoxyheptane (**24**) and *trans-1*-phenylpropylene oxide (**25**) were synthesised mixing the respective alkene (**3**, **4** and **5**) to a final concentration of 200 mM in water with 30% acetonitrile and 10% of acetone. 3 eq. of oxone® was added portion wise over 4 hours. The pH was kept at 8 by adding $NaHCO_3$ to the solution.[153] The reactions were followed by TLC and when no more substrate was visible the reaction mixture was extracted three times with *tert*-

butylmethylether (TBME). The combined organic layers were washed with brine (1×), dried over Na_2SO_4 and filtered before solvent removal by evaporation. The epoxides were purified using column chromatography with cyclohexane/isopropanol (95:5) for **23**, cyclohexane/ethyl/acetate (95:5) for **24** and (90:10) for **25**. The synthesised epoxides were confirmed using NMR (300 MHz; diluted solution in $CDCl_3$) and literature data.[154–156]

Chemical synthesis of authentic azido alcohol standards

1 mmol of either *trans*-2,3-epoxyhexane (**21**) (100 mg) or *trans*-2,3-epoxyheptane (**22**) (114 mg) were mixed with 202 mg of NaN_3 (3.1 eq.) and 166 mg of NH_4Cl (3.1 eq.) in 3.5 mL of methanol and refluxed until no more substrate was visible on TLC (5-6 h). The reaction mixture was then diluted in diethyl ether, washed with brine and the water phase was extracted twice with diethyl ether. The organic layers were combined and dried over Na_2SO_4. The crude extracts were purified by column chromatography with cyclohexane/diethyl ether (90:10) yielding 52% azido-hydroxyhexane (**31**) and 48% azido-hydroxyheptane (**32**).[157]

Following a similar protocol the azido alcohols **33**, **34** were synthesised by dissolving the respective epoxides in MeOH with 5% H_2O to a final concentration of 200 mM, NaN_3 (3.1 eq.) and NH_4Cl (3.1 eq.) was added and the reaction was stirred overnight at reflux. Methanol was evaporated, the crude extract was dissolved in *tert*-butylmethylether (TBME) and washed with the same volume of brine (1x) and the water phase was extracted three times with TBME. The combined organic layers were dried over Na_2SO_4 and filtered before solvent removal by evaporation. **33** and **34** were purified by column chromatography with chloroform/acetone (95:5) and cyclohexane/ethyl acetate (95:5), respectively.

Chemical synthesis of authentic amino alcohol standards

The azido group hydrogenation protocol was adapted from a previously published method.[158] 200 mg of pure azido alcohol **31** or **32** (~60 mM) were dissolved in 20 mL methanol. Pd/C in larger quantities (40 mg) is pyrogenic. Therefore, it was pre-wetted with methanol before adding it to the mixture. The solution was magnetically stirred at 500 rpm at room temperature and hydrazine was added dropwise for 2 h until a concentration of 150 mM was reached (2.5 equivalents). After ~4 h the reaction was filtered through Celite® to separate Pd/C and the solvent was removed with vacuum. Amino alcohols **41** and **42** were purified by column chromatography using a mixture of chloroform:methanol:ammonia (90:9:1).

2.1.2 Kits, equipment and software

Table 3. Kits used in this study.

Kit	Supplier
NucleoSpin® Gel and PCR Clean-up Kit	Macherey-Nagel
NucleoSpin® Plasmid Purification Kit	Macherey-Nagel
QIAquick® Gel Extraction Kit	Quiagen
Heterogeneous Palladium catalyst Kit I	Sigma-Aldrich

Table 4. Equipment used in this study.

Equipment	Supplier
ÄKTA pure	GE Healthcare
HisTrapTM FF column	Macherey-Nagel
CARY 60 Bio UV-Vis spectrophotometer	Agilent Technologies
Helios UV / Vis Spectrophotometer	Unicam
Ultrospec 3000 UV/Visible Spectrophotometer	Pharmacia Biotech
Implen NanoPhotometer NP80 Touch	Intas-Science-Imaging Instruments GmbH
NanoDrop™ 1000 spectrophotometer	peqlab
Heraeus Fresco 21 Centrifuge	Thermo Fisher Scientific
Heraeus Fresco 17 Centrifuge	Thermo Fisher Scientific
Heraeus Pico 17 Centrifuge	Thermo Fisher Scientific
Heraeus Primo R Centrifuge	Thermo Fisher Scientific
Heraeus Multifuge X3R Centrifuge	Thermo Fisher Scientific
Mikro 185 Centrifuge	Andreas Hettich GmbH & Co.KG
PeqSTAR Thermocycler	VWR Peqlab
Gel Doc™ XR+ System	Biorad
ThermoMixer C	Eppendorf
Microx 4	PreSens Precision Sensing GmbH
Profiling oxygen microsensor (needle type)	PreSens Precision Sensing GmbH
Vibra-Cell™ VCX130	Sonics & Materials, Inc.
Bioflo 110	New Brunswick Scientific
Laborota 4001	Heidolph Instruments GmbH & Co Kg
SC 920	KNF Neuberger GmbH
Speed Vac® SC 110	Savant
Gel Doc™ XR+ System	Biorad
GC2010 plus gas chromatograph with AOC-20i-injector and FID detector	Shimadzu
GC2014 gas chromatograph with AOC-20i-injector and FID detector	Shimadzu
Supreme 5ms column	CS Chromatographie Service
HYDRODEX γ-DiMOM column	Macherey-Nagel
Lipodex® E column	Macherey-Nagel
CP-Sil 5 CB	Agilent Technologies, Inc
BPX5 column	SGE Analytical Science
HYDRODEX β-3P column	Macherey-Nagel
Lipodex® G column	Macherey-Nagel

Table 5. Software used in this study.

Software	Supplier
PyMOL	DeLano Scientific LLC.
Rosetta 3.6	RosettaCommons
UNICORN (5.31)	GE Healthcare
ChemDraw Professional 15.1	Perkin Elmer
Clone Manager (9.1)	Scientific & Educational Software

Serial Cloner 2.6	SerialBasic
GCsolution (2.32.00)	Shimadzu
GCMSsolutions (2.53)	Shimadzu
LabSolution (5.51)	Shimadzu
MestRenova 6.0	Mestrelab Research
Origin 2017	OriginLab
TopSpin 4.0.3	Bruker BioSpin GmbH
PreSens measurements Studio 2	PreSens Precision Sensing GmbH
RStudio 1.0.136	RStudio, Inc.

2.1.3 Bacterial strains, plasmids and oligonucleotides

The *Escherichia coli* DH5α strain (Invitrogen, Carlsbad, CA, USA) was used for plasmid amplification and DNA manipulation, while *E. coli* BL21 (DE3) (Novagen, EMD Biosciences, San Diego, CA, USA) and *E. coli* C43 (DE3) (Lucigen, Middleton, WI, USA) were used for heterologous expressions.

The vector pMK-RQ containing the codon-optimized genes coding for styA and styB was ordered from Invitrogen (Thermo Fisher Scientific division, Wilmington, USA).

The alcohol dehydrogenase from *Lactobacillus kefir* (LkADH) was present in pET21(+) vector and was used as it is from Enzymicals' cofactor regeneration enzyme collections.

The alcohol dehydrogenase from *Leisfonia sp.* S749 (LsADH) was ordered cloned in pCDF1Duet-1, while the gene coding for StyB was ordered cloned in pRSFDuet-1 (Genescript, New Jersey, U.S.).

The vectors pET28a(+)-cyp154C5, pET28a(+)-cyp154H1, pACYC-Duet1-camA-camB, pACYC-camA, pACYC-camB were readily available from previous projects.[159,160]

The vector pET28a(+)-MsAct was present in the plasmid collection from a previous project.[161]

All HHDH genes except *hheA2*, *hheB2* and *hheC* were readily available in pET28a(+)-based vectors.[102] Instead HheA2, HheB2 and HheC genes were present in pBAD vector, as described previously.[24,162]

styAB subcloning in pETDuet and pRSFDuet-1 vector

Synthetic styB gene was cut from the commercial plasmid using the endonucleases XhoI and NdeI (NEB). In parallel pETDuet vector was digested using the standard NEB protocol with the same restriction enzymes. Afterwards, the cut plasmid and styB were excised from 0.8% agarose gel and purified via the QIAquick® Gel Extraction Kit (Qiagen, Hilden, Germany). The ligation was carried out overnight at 26 °C using the T4 ligase and a molar ratio of 1:3 (vector to insert). After transformation of chemically

competent *E.coli* DH5α, the successful insertion in the MSC-2 was verified by colony PCR using the protocol described below (Section 2.2.1).

The gene coding for styA was directly adjacent to styB with only one restriction site in between. For viable cloning in the MSC-1 the restriction sites at both ends were modified via PCR using primers with modified ends. The 5'-end of styA was modified using Fw_StyA_NcoI, while the 3'-end was modified by using the primer Rv_StyA_BamHI. Hence, the restriction sites were exchanged from NdeI to NcoI at the 5´-end and from HindIII to BamHI at the 3'-end. The gene was amplified by PCR as described in section 2.2.1 and was digested at 37 °C using NcoI and BamHI and purified using the NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel). In parallel, pETDuet-styB was digested with the same endonucleases and excised from the agarose gel as described above. The ligation was performed overnight using the standard protocol and a molar ratio of 1:3 (vector to insert). The successful insertion in the MSC-1 was then verified after transformation of chemically competent *E.coli* DH5α by colony PCR using the protocol described below. Finally, plasmid DNA was isolated with NucleoSpin Plasmid kit (Macherey Nagel, Düren, Germany) and digested using the endonucleases XhoI, NdeI, NcoI and BamHI to verify the presence of the two inserts. Additionally, an aliquot was sent for sequencing to Eurofins Genomics.

The cloning of styA from pETDuet-styAB to pRSFDuet-1-styB was performed as described above. pETDuet-styAB and pRSFDuet-1-styB were digested at 37 °C using NcoI and HindIII. Open vector and gene of interest were purified using the QIAquick® Gel Extraction Kit (Qiagen, Hilden, Germany). The ligation was performed overnight using the standard protocol for T4 ligase and a ratio of 1:3 (vector to insert). The successful insertion in the MSC-1 was then verified by sequencing of the isolated plasmid DNA after transformation of chemically competent *E.coli* DH5α plasmid extraction and sequencing using GATC Biotech sequencing (Konstanz, Germany).

Table 6. List of primers used for gene amplification and sequencing and relative lengths and melting temperatures

Oligo's name	Sequence (5' → 3')	Length (bp)	T_m (°C)
Fw_StyA_NcoI	GGAGAC**CC**ATG**G**CCAAGAGCATCGCCATCGTC	32	58.6
Rv_StyA_BamHI	AGGCCC**GGATCC**TTACACACCCGTACCTGCGAAC	34	59.3
T7-fwd	TAATACGACTCACTATAGGG	20	53.2
T7-rev	GCTAGTTATTGCTCAGCGG	19	56.7
T7 term	CTA GTT ATT GCT CAG CGG T	19	54.5
ACYCDuetUP1	GGATCTCGACGCTCTCCCT	19	61.0
DuetDOWN1	GATTATGCGGCCGTGTACAA	20	57.3
petup	ATGCGTCCGGCGTAGA	16	54.3

2.2 Methods

2.2.1 Molecular and microbiological methods

Cultivation and conservation of *Escherichia coli*

Escherichia coli cells were cultivated in LB (5 g/L peptone, 10 g/L NaCl, 10 g/L yeast extract) or TB (4 mL/L glycerol, 12 g/L peptone, 24 g/L yeast extract, 0.17 M KH_2PO_4, 0.74 M K_2HPO_4) medium supplemented with the respective antibiotics (Table 7) at 37 °C overnight. The antibiotic was added after sterilization of the medium by autoclave. Agar plates were prepared by adding agar-agar (1.2% w/v) to the medium before autoclaving.

For long-term conservation of *E. coli* cells, 900 µL of an overnight culture were mixed with either 60 µL of DMSO or alternatively 500 µL of *E. coli* cells were mixed with 500 µL of a solution of 60% v/v glycerol and stored at -80 °C.

Table 7. List of vectors and respective antibiotic used during this work. The strains used are described in section 2.3.1

Vectors	Antibiotics
pMK-RQ-styAB	Kanamycin (50 µg/mL)
pETDuet-styB or pETDuet-styAB,	Ampicillin (100 µg/mL)
pRSFDuet-1-styB or pRSFDuet-1-styAB	Kanamycin (50 µg/mL)
pETDuet-styAB, pG-KJE8	Ampicillin (100 µg/mL), Chloramphenicol (20 µg/ml), Tetracycline (5 ng/m), Arabinose (0.5 mg/m)
pET21(+)-LkADH	Ampicillin (100 µg/mL)
pCDF1Duet-1-LsADH	Spectinomycin (100 µg/mL)
pRSFDuet-1-styAB, pET21(+)-LkADH	Kanamycin (50 µg/mL), Ampicillin (100 µg/mL)
pRSFDuet-1-styAB, pCDF1Duet-1-LsADH	Kanamycin (50 µg/mL), Spectinomycin (100 µg/mL)
pETDuet-styAB, pCDF1Duet-1-LsADH	Ampicillin (100 µg/mL), Spectinomycin (100 µg/mL)
pET28a(+)-cyp154C5 or pET28a(+)-cyp154H1	Kanamycin (50 µg/mL)
pACYC-camA or pACYC-camB or pACYC-Duet1-camA-camB	Chloramphenicol (25 µg/ml)
pET28a(+)-cyp154C5, pACYC-Duet1-camA-camB	Kanamycin (50 µg/mL), Chloramphenicol (25 µg/ml)
pET28a(+)-cyp154H1, pACYC-Duet1-camA-camB	Kanamycin (50 µg/mL), Chloramphenicol (25 µg/ml)
pET28a(+)-MsAct	Kanamycin (50 µg/mL)
pET28a(+)-hhdh except hheA2, hheB2 and hheC	Kanamycin (50 µg/mL)
pBAD-hheA2 or pBAD-hheB2 or pBAD-hheC	Ampicillin (100 µg/mL)

Preparation of competent *E. coli* cells

Depending on the genetic transformation efficiency wanted two different types of competent cells, chemically competent or electrocompetent, were used in this study.

Chemically competent cells preparation

For the preparation of chemically competent cells of *E. coli* DH5α, Top10, BL21 (DE3) and C43 (DE3) 50 mL LB medium was inoculated with an overnight culture of the desired bacterium to an optical density at 600 nm (OD_{600}) of 0.1. The culture was incubated at 37 °C at 250 rpm until an OD_{600} between 0.4 and 0.6 was reached. After incubation on ice for 15 min cells were harvested by centrifugation (4000 rpm, 4 °C, and 30 min). The cell pellet was resuspended in 20 mL ice-cold RF1 buffer (30 mM KCH_3CO_2, 50 mM $MnCl_2$, 100 mM $RbCl_2$, 10 mM $CaCl_2$, 15%(v/v) glycerol, pH 5.8) and incubated for 2 h on ice. Afterwards, the mixture was centrifuged and the resulting pellet was resuspended in 4 mL RF2 buffer (10 mM MOPS, 75 mM $CaCl_2$, 10 mM $RbCl_2$, 15% (v/v) glycerol, pH 6.8) and in 200 µL aliquoted. Aliquots were shock frozen in liquid nitrogen and stored at -80 °C.[163]

Electrocompetent cells preparation

E. coli cells were cultivated and harvested as described above. The cell pellet was washed 3 times with 200 mL ice-cold sterile ddH_2O followed by washing with ice-cold 20% (v/v) glycerol/water solution and centrifugation. Finally, the cell pellet was resuspended in 2 mL ice-cold 10% (v/v) glycerol solution and 50 µL aliquots were prepared. Aliquots were shock frozen in liquid nitrogen and stored at -80 °C.[163]

Transformation of *E. coli*

Chemically competent *E. coli* cells were transformed by adding 50 to 100 ng of plasmid DNA to a cell aliquot. The cells were incubated for 20 min on ice and heat shock was performed for 90 s at 42°C. 600 µL of SOC medium (2 g/L KCl, 2 g/L NaCl, 5 g/L yeast extract, 10 g/L tryptone, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was added and cells were incubated for 1 h at 37 °C at 800 rpm for cell recovery and antibiotic resistance expression.

Electrocompetent *E. coli* cells were transformed by adding 50 to 100 ng of plasmid DNA. The cells were thawed on ice, gently mixed and transferred in a pre-cooled 2 mm cuvette for electroporation. The electroshock was carried out using Electroporator 2510 (Eppendorf AG, Hamburg, Germany) at 2500 V and 5 ms. 600 µL of SOC medium was readily added to the cuvette and the mixture was transferred to a 1.5 mL reaction tube to incubate for 1 h at 37°C at 800 rpm.

At last, the transformed cells were plated on LB agar plates containing the desired selection marker (antibiotic, Table 7).

Standard protocol for agarose gel electrophoresis

Agarose gel electrophoresis was generally used for plasmid visualization and vector insert separation after restriction enzyme's digestion. Agarose (0.8 or 1% w/v) was dissolved in TAE buffer (40 mM Tris-HCl, 2 mM acetic acid, 1 mM EDTA, pH 8.5) and the gel was formed via microwave oven. DNA samples were mixed with 6x loading dye (New England Biolabs, Frankfurt, Germany) and typically 10-20 µL were loaded on the gel together with 1 kb ladder (NEB) to reference the size of the DNA fragments. The separation based on molecular weight using an electric current of 100 V. TAE buffer was used as running buffer as well. The gels were stained for 15 in a water bath containing ethidium bromide and visualized under UV light. The DNA concentration of samples was determined using a NanoDrop™ 1000 spectrophotometer.[164]

Polymerase chain reaction (PCR)

The polymerase chain reaction is an extremely versatile laboratory technique that allows amplifying a selected DNA sequence.[165,166] Depending on the oligonucleotides (primers) used to drive the amplification different results can be obtained. In this work, the PCR was applied to selectively modify the restriction sites at both ends of the gene coding for StyA, to verify if the ligation was successful (colony PCR) or to introduce mutations in hheG gene via site-directed mutagenesis.

Gene of interest amplification and selective restriction site modification

20-100 ng template DNA was added to a solution containing 1 U Phusion© polymerase (New England Biolabs, Frankfurt, Germany), 5% (v/v) 5x Phusion HF buffer, 200 µM dNTPs. The target DNA sequence is then selectively amplified using designed forward and reverse primers to a final concentration of 0.5 µM, finally the volume was adjusted to 50 µL with ddH_2O. The gene of interest was amplified according to the temperature profile in Table 8.

Table 8. Settings used in the thermocycler for the PCR.

Step	Temperature (°C)	Time (min)	Cycles
1	98	30 s	1
2	98	10 s	
3	Tm primer -5°C	30 s	25 - 35
4	72	30 s per kb	
5	72	10 min	1

The DNA fragments of interest were separated via agarose gel electrophoresis, excised from agarose gels and purified using QIAquick® Gel Extraction Kit prior to further experiments.

Colony PCR

Colony PCR is a powerful tool to quickly determine if a ligation was successful. After *E. coli* DH5α was transformed with the ligation product, individual colonies were picked from a LB-agar plate, dissolved in 20 µL ddH_2O and transferred on a fresh LB-agar plate. The mixture was heated for 2 min at 98 °C and 2 µL of the cell suspension were used for the amplification. The sample was mixed with DreamTaq Green DNA Polymerase (0.5 µL), 10X DreamTaq Green Buffer, 5 µM forward and reverse primer and ddH2O to a final volume of 50 µL. The gene of interest was amplified according to the temperature profile in Table 9. After amplification, the mix was directly loaded on the agarose gel and the results directly visualized after electrophoresis.

Table 9. Temperature profile used in the thermocycler for colony PCR.

Step	Temperature (°C)	Time (min)	Cycles
1	94	5	1
2	94	1	
3	primer Tm-5°C	0.5	30
4	72	1	
5	72	3	1

QuickChange PCR method

The site-directed mutagenesis was performed by Janine Mayer (Master thesis, 2017) to analyse the mutations suggested by the *in silico* mutagenesis done with the coupled moves method (Section 3.3.3). 100 ng of parental vector DNA and 125 ng of the respective forward and reverse primer were added to a 25 µL aliquot of the PCR Mastermix and the volume was adjusted to 50 µL with ddH_2O. The temperature profile used for the amplification is shown in Table 10.

Table 10. Temperature profile used in the thermocycler for the QuickChange PCR protocol.

Step	Temperature (°C)	Time (min)	Cycles
1	95	2	1
2	95	0.5	
3	55	1	30
4	68	1 min/kb	

The template vector DNA was digested overnight using 1 µL of *Dpn*I and incubation at 37 °C. Afterwards, chemically competent *E. coli* DH5α cells were transformed with 5 µL of the PCR mixture and plated on LB agar plates The plasmid DNA was isolated, from 5 mL of liquid culture, the amino acid exchange was confirmed by sequencing (Eurofins Genomics). More rounds were performed to generate multiple mutations in the same vector.

2.2.2 Biochemical methods

Recombinant protein expression in *E. coli*

Flask cultivations for protein expression

Heterologous protein expression was generally done in 400 mL TB media (2 L flask) with the respective selection markers shown in Table 7. A colony from the desired LB-agar place was picked to inoculate 5 mL of LB media for the overnight culture. The ON-culture was then used to inoculate 50 ml TB media pre-culture (250 mL flask) to allow the cells to adapt to the new medium. The cells were grown at 37 °C at 200 rpm until they reached OD_{600} of approx. 2-4. The main culture (400 mL) was then inoculated with a volume to have a starting OD_{600} of 0.1. The cells were grown at 37 °C at 200 rpm until they reached OD_{600} of 0.8-1.2, afterwards the desired amount of Isopropyl-β-D-thiogalactopyranoside (IPTG) was supplemented as inducer and the temperature was lowered to the required expression temperature. The induction was performed overnight except whereas protein production in *E. coli* C43 (DE3) required 48 h. Optimal temperatures and IPTG concentrations for the different combinations of strain and expression vector used are shown in Table 11. After the expression time, the cells were harvested by centrifugation at 4000 rpm for 30 min at 4 °C, cell pellets were washed and stored at -20 °C until further experiments. Alternatively, cells pellets were resuspended in 50 mM potassium phosphate buffer pH 7.5 to a concentration of 0.60 g/mL and cell disruption was performed via sonication (6 cycles of 30 s at an amplitude of 50%, with 30 s pause on ice between cycles). Cellular debris was removed by centrifuging the cellular lysate at 10,000 rpm for 30 min at 4 °C. Glycerol to a final concentration of 20% was added to the cell-free extract to a final concentration of 75 mg/ml ca. and stored at -20 °C.

Table 11. Expression conditions used for the different strains and heterologous proteins. *Additional 0.5 mM of δ-amino levulinic acid (δ-ala) and 0.1 % v/v of trace element solution (0.50 g/L of $CaCl_2 \cdot 2H_2O$, 0.18 g/L of $ZnSO_4 \cdot 7H_2O$, 0.10 g/L of $MnSO_4 \cdot H_2O$, 20.10 g/L of Na_2-EDTA, 16.70 g/L of $FeCl_3 \cdot 6H_2O$, 0.16 g/L of $CuSO_4 \cdot 5H_2O$ and 0.18 g/L of $CoCl_2.6H_2O$)

Strain	Vectors	Temperature (°C)	IPTG (mM)
E. coli BL21 (DE3)	pETDuet-styAB,	20 °C	0.2 or 0.5 mM
E. coli C43 (DE3)	pETDuet-styAB,	25 or 30 °C	0.2 or 0.5 mM
E. coli BL21 (DE3)	pETDuet-styAB, pG-KJE8	20 °C	0.2 or 0.5 mM
E. coli C43 (DE3)	pETDuet-styAB, pG-KJE8	25 or 30 °C	0.2 or 0.5 mM
E. coli BL21 (DE3)	pRSFDuet-1-styAB	20 °C	0.2 mM
E. coli BL21 (DE3)	pET21(+)-LkADH	20 °C	0.2 mM
E. coli BL21 (DE3)	pCDF1Duet-1-LsADH	20 °C	0.2 mM
E. coli BL21 (DE3)	pRSFDuet-1-styAB, pET21(+)-LkADH	20 °C	0.2 mM
E. coli BL21 (DE3)	pRSFDuet-1-styAB, pCDF1Duet-1-LsADH	20 °C	0.2 mM
E. coli BL21 (DE3)	pETDuet-styAB, pCDF1Duet-1-LsADH	20 °C	0.2 mM
E. coli C43 (DE3)	pET28a(+)-cyp154C5 or pET28a(+)-cyp154H1	30 °C	0.8 mM*

E. coli C43 (DE3)	pACYC-camA or pACYC_camB or pACYC-Duet1-camA-camB	30 °C	0.8 mM*
E. coli C43 (DE3)	pET28a(+)-cyp154C5, pACYC-Duet1-camA-camB	30 °C	0.8 mM*
E. coli C43 (DE3)	pET28a(+)-cyp154H1, pACYC-Duet1-camA-camB	30 °C	0.8 mM*
E. coli BL21 (DE3)	pET28a(+)-MsAct	20 °C	0.2 mM
E. coli BL21 (DE3)	pET28a(+)-HheE	20 °C	0.2 mM
E. coli BL21 (DE3)	pET28a(+)-HheE5	20 °C	0.2 mM
E. coli BL21 (DE3)	pET28a(+)-hhdh	Previous results[167]	
E. coli TOP10	pBAD-hhdh	Previous results[24,162]	

Cultivation of *E. coli* at 2 L bioreactor scale

Production of the biocatalyst selected for the cascade was also tested at a larger scale. *E. coli* BL21 (DE3) cells harbouring either pETDuet-styAB and pCDFDuet-1-LsADH or pET28a(+)-HheE or pET28a(+)-HheE5 were grown overnight in 5 mL LB medium with the appropriate selection marker at 37 °C and 150 rpm. To ensure enough biomass for inoculation at the larger scale a 200 mL pre-culture in TB medium was inoculated with the 5 mL overnight culture. When the OD_{600} reached ~2-3, finally the 2 L bioreactor (Bioflow 110, New Brunswick Scientific) containing 1.8 L of TB medium was inoculated taking care to adjust the starting OD_{600} to 0.1. The pH was set to 7.0 and titrated during the cellular growth with either 5 M KOH or 1 M H_3PO_4 solutions. Compressed air was provided with a 1.5 L/min flow and the dissolved oxygen was controlled via pO_2 sensor and set to be at least 30%. The stirring speed was set to be variable depending on the dissolved oxygen values from 300 to 800 rpm. Finally, excessive foam formation was controlled with the commercial Antifoam 204 from Sigma Aldrich Chemie (Steinheim, Germany). Protein expression was started when the OD_{600} reached a value of approx. 0.8-1.2 with 0.2 mM IPTG as inducer and the temperature was lowered to 20 °C. Samples at different time points were taken to measure the cellular density and enzyme activity (for StyAB) and after 16-22 h the cells were harvested by centrifugation (4000 rpm, 30 min, and 4 °C). Cell pellets were washed with 50 mM potassium phosphate buffer pH 7.5, the cellular suspension was centrifuged again and pellets were stored at -20 °C until further use.

SDS-PAGE

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is a commonly used technique that allows protein separation by applying an electric field. SDS is used to denaturate and to homogenize the number of charges around the protein resulting in separation based on the molecular weight.[168] For the typical separation 12% (v/v) acrylamide gels were used (composition in Table 12). To assess the protein expression, the cellular suspensions were normalized by adjusting the concentration to

OD_{600} ~10 in 1 mL in 50 mM potassium phosphate buffer pH 7.5. The samples were lysed by sonication (2 cycles of 30 seconds pulse with each 45 seconds pause, 50% of amplitude). After centrifugation at 13.300 rpm for 10 minutes, 30 µL of supernatant were mixed with 10 µL of loading dye (50 mM Tris HCl pH 6.8, 10 % (v/v) glycerol, 4 % (w/v) SDS, 0.03 % (w/v) bromophenol blue and 15 mg/mL DTT) and incubated at 95°C for 5 min. The remaining pellet was resuspended in 1 mL of 50 mM potassium phosphate buffer pH 7.5. Afterwards, 30 µL of the solution were mixed with 10 µL of loading dye and treated as mentioned previously. In all cases, 12 µL samples were loaded and gels were run at 80 V for 10 min to allow the protein sample to align during the stacking gel, followed by another 60 min ca. at 120 V for separation.

Table 12. Composition of the two gels that composes a 12% Tris-glycine SDS-PAGE gel

Component	Stacking gel (6% acrylamide)	Resolving gel (12% acrylamide)
ddH2O	1.5 mL	2.25 mL
Acrylamide (40%)	375 µL	1.5 mL
1,5 M Tris-HCl (pH 8.8) 0.4 % SDS	-	1.25 mL
1 M Tris-HCl (pH 6.8) 0.4 % SDS	625 µL	-
10 % APS	25 µL	50 µL
TEMED	2.5 µL	5 µL

Purification of His-tagged proteins

The expression of proteins with a C or N-terminal hexahistidine tag from the pET28a(+) based expression vector system, allows the facile protein purification by immobilized metal affinity chromatography (IMAC) using a 5 mL HisTrap HP column and Äkta FPLC system (GE Healthcare, Freiburg, Germany). Cells pellets, prepared as described previously, were resuspended in the equilibration buffer (50 mM Tris·SO_4, 500 mM Na_2SO_4, 25 mM imidazole, pH 7.9) containing 1 mg/mL lysozyme and 100 µM phenylmethylsulfonylfluoride (PMSF) as protease inhibitor to a final concentration of 0.50 mg/mL. Cell disruption was performed by sonication (6 cycles of 30 s at an amplitude of 50%, with 30 s pause on ice between cycles) and cell-free extract (CFE) was produced by centrifuging the cellular lysate at 10000 rpm for 30 min at 4 °C. Further, CFE was filtered through a 0.45 µm cellulose acetate membrane filter prior loading it on an equilibrated column (5x column volumes (cv) of equilibration buffer). The filtered CFE was loaded with a flow rate of 2 mL/min and the non-specifically bound proteins were eluted by washing the column with 5 cv of equilibration buffer with a flow rate of 5 mL/min. The His-tagged protein was eluted with an imidazole concentration gradient from 25 to 500 mM realized by mixing the equilibration buffer with elution buffer (50 mM Tris·SO_4, 500 mM Na_2SO_4, 500 mM imidazole, pH 7.9) in different percentages over 100 mL with a flow rate of 2 mL. Fractions of 2 mL were collected and the elution of the target protein

was followed by the intensity of the UV signal at 280 nm. 30 µL samples of elution fractions were analysed by SDS-PAGE to determine the purity. Fractions with lowest impurity content were pooled together and concentrated via ultrafiltration (Amicon Ultra-15 Centrifugal Filter with 10 kDa cut-off, Merck Millipore, Darmstadt, Germany). PD-10-columns (GE Healthcare) were used to change the buffer from elution to storage buffer (TE buffer: 10 mM Tris·SO4, pH 7.9, 4 mM EDTA) containing 10% (v/v) glycerol. The protein concentration was calculated with the extinction coefficient of the purified protein and the absorption at 280 nm measured with the NanoDrop™ 1000 Spectrophotometer. 100 µL aliquots were stored at -20 °C until further use. Cytochrome P450 monooxygenases concentration was determined by CO-difference spectroscopy calculating the maximum absorbance of CO-bound at 450 nm (ε_{450}=91 mM^{-1} cm^{-1}).[169]

Purification of non-his-tagged proteins

PdR and Pdx from *P. putida* were both expressed without an affinity tag and therefore the two proteins were purified in two steps.

Ion exchange chromatography

The first step was in common for both proteins and the cell-free extracts were prepared as described before, with the exception that the equilibration buffer was different (50 mM potassium phosphate buffer pH 7.5). Using a 120 mL Q-Sepharose FF ion exchange column equilibrated with the equilibration buffer PdR or Pdx were loaded with a flow of 6 mL/min. Proteins were eluted using a 1 L gradient from 0 to 500 mM KCl in 50 mM potassium phosphate buffer pH 7.5 with a flow of 6 mL/min. PdR and Pdx eluted at approx. 350 and 300 mM of the gradient, respectively. It was possible to follow the separation of PdR and Pdx by their colour as PdR has a typical yellow colour given by FAD cofactor. Whereas Pdx possesses the typical brown colour of proteins containing an iron-sulphur cluster. Elution fractions containing PdR (45 kDa) were concentrated by ultrafiltration (Amicon Ultra-15 Centrifugal Filter with 30 kDa cut-off, Merck Millipore, Darmstadt, Germany) and submitted to the next purification step. The fractions containing Pdx (11 kDa) were pooled together and further purified by ultrafiltration (Amicon Ultra-15 Centrifugal Filter with 30 kDa cut-off, Merck Millipore, Darmstadt, Germany). Hence, larger proteins can be removed and the flow through containing purified Pdx was concentrated using a 10 kDa cut-off Amicon centrifugal filter. Desalting and storage were carried out as described previously.

Hydrophobic interaction chromatography

The concentrated PdR fractions obtained by ion exchange chromatography were dissolved in 20 mL 50 mM potassium phosphate buffer pH 7.5 containing 1.5 M KCl and loaded on a 50 mL Phenyl-Sepharose FF hydrophobic interaction column equilibrated

with the same buffer. The elution was started with a gradient from 1.5 to 0 M of KCl in 500 mL with a linear flow of 6 mL/min. PdR eluted approximately at 1.12 M KCl (15% of the gradient), the best purity fractions containing PdR were pooled together concentrated, desalted and stored as described previously. The concentration of pure PdR was determined at 454 nm (ε_{455}=10.0 mM^{-1} cm^{-1}) while the concentration of Pdx at 455 nm (ε_{455}=10.4 mM^{-1} cm^{-1}). The electron transfer components (ETC) Pdx and PdR (ration 3:16) was determined by monitoring the increase in absorbance at 550 nm of reduced cytochrome c (ε_{450}=19.1 mM^{-1} cm^{-1}).[170]

Enzyme assays

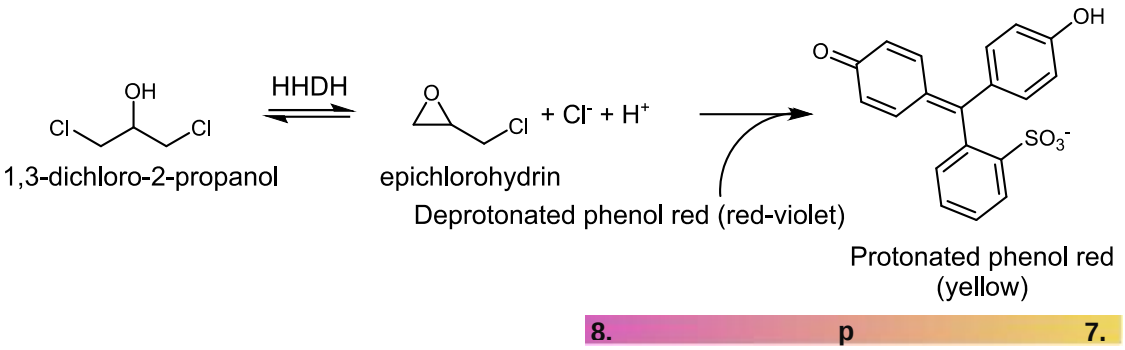
Styrene monooxygenase

The activity of whole cells or cell-free extracts harbouring the styrene monooxygenase was measured via *trans*-2,3-epoxyheptane (**22**) formation from the substrate *trans*-2-heptene (2.5 mM) (**2**) by adjusting a previously published protocol.[78] Whole cells (60 g/L wet cell weight) or cell-free extract (2.3 g/L) were resuspended in 50 mM potassium phosphate buffer pH 7.5 to a final volume of 1 mL. In the case of whole cells, glucose was added to a final concentration of 10 mM and the cells were incubated at 30 °C, with magnetic stirring at 400 rpm for 3 min to allow the metabolism to start. While cell-free extract typically contained LsADH and therefore the cofactor regeneration was performed using the isopropanol derived from the substrate stock solution. In both cases substrate (**2**) was added to a final concentration of 3.5 mM from a 50 mM stock solution in isopropanol and the reactions were incubated at 30 °C, with magnetic stirring at 400 rpm for 5 min. After 5 min reaction, a 400 µL sample was transferred to a 1.5 mL reaction tube stored on ice. Immediately, 400 µL of TBME containing 0.1% (v/v) dodecane as an internal standard was added and the sample was mixed using a vortex for 1 min. After phase separation by centrifugation the organic phase was dried over anhydrous sodium sulphate and submitted to gas chromatography for analysis. One unit of enzyme activity was determined as product formation with a rate of µmol/min. Each experiment was carried out in duplicate.

Halohydrin dehalogenases

The activity of halohydrin dehalogenases was determined directly using cell-free extract with a colourimetric assay measuring the pH variation caused by the dehalogenation reaction carried out by the enzymes (Scheme 16). The cell-free extract solution was prepared in HEN buffer (2 mM Hepes, 0.5 mM EDTA, 10 mM, Na_2SO_4, pH 8.2) to a final total protein content of 10 mg/ml and the pH was carefully adjusted to 8.2. Typically, 15 µL of the phenol red solution (1 mg/mL in 20% ethanol) were dissolved in 925 µL HEN buffer and incubated at 30 °C and 600 rpm for 5 min. Afterwards, 10 µL of the substrate

solution (1M 1,3-dichloro-2-propanol in DMSO) were added to the solution and 50 µL of the enzyme solution were added. The cuvette was immediately inserted in the spectrophotometer and the decrease in absorbance at 560 nm was monitored for 1 min. 1 U of enzyme activity is defined as the amount of enzyme necessary to generate 1 µmol/min of H^+ from dehalogenation at 30 °C. The change of concentration of H^+ is calculated using a calibration curve where HCl (10 mM stock solution in ddH_2O) is added at different concentrations (0-1 mM). The calibration curve must be rejected if the square of the Pearson coefficient (R^2) is lower than 0.98. All experiments were performed at least in triplicate due to the high standard deviation of this assay.



Scheme 16. Reaction scheme of the colourimetric assay used to measure the HHDH activity.

Enzyme kinetics

The kinetics parameters for HheE and HheE5 were determined by fitting the experimental data with either the Michaelis-Menten equation (1)[171] or the Hill's equation (2) to include cooperative bindings in the fitting.[172,173] Reaction rates (V) at different substrate or NaN_3 concentrations ([S]) were measured in order to determine K_M/K_{50}(substrate concentration at which the reaction rate is at half-maximum), V_{max} (maximum rate at a given enzyme concentration).

$$V = \frac{V_{max} \cdot [S]}{K_m + [S]} \qquad (1) \qquad V = \frac{V_{max} \cdot [S]^n}{(K_{50})^n + [S]^n} \qquad (2)$$

The kinetic parameters relative to HheE and *trans*-2,3-epoxyhexane (**1**) were calculated by keeping the enzyme and sodium azide concentrations constant and varying the substrate concentration. Initial reaction rates were calculated at 5, 10, 25, 50, 100, 150 and 200 mM of **1**, keeping HheE concentration at 300 µg/mL and NaN_3 at 50 mM.

The kinetic parameters relative to HheE and *trans*-2,3-epoxyheptane (**2**) were calculated as described above using different substrate concentrations (10, 20, 50, 100, 150 and 200 mM).

To determine the kinetic values for sodium azide for HheE, *trans*-2,3-epoxyhexane (**1**) concentration was kept constant at 50 mM and the reaction rate was calculated at 5, 10, 20, 50 and 150 mM of NaN3.

In a similar way, HheE5's kinetic data was obtained for *trans*-2,3-epoxyheptane (**2**) where the concentration was varied from 5,10, 20, 50, 80, 100 to 120 mM. As described above for HheE, HheE5 concentration was kept at 150 µg/mL and NaN_3 concentration at 50 mM.
Similarly, the kinetic values for sodium azide were determined by keeping constant *trans*-2,3-epoxyheptane (**2**) concentration at 50 mM and varying sodium azide concentration (10, 20, 50, 100, 120, 150 and 200 mM).

All reactions were performed in 1.2 mL at 25 °C and samples were taken at 60, 65, 75 and 90 min as no product was formed in the first hour. The samples were prepared and analysed in GC using the protocol described below.

2.2.3 Analytical methods

Sample preparation

Typically an extraction with an organic solvent is required when reactions are performed in aqueous buffer as GC columns are very sensible to the presence of water. For this, generally a 400 µL sample from the aqueous phase was mixed with 400 µL of TBME or ethyl acetate containing 0.1% (v/v) dodecane as an internal standard and mixed by vortexing for 1 min. The organic phase was separated by centrifugation, dried over anhydrous sodium sulphate or magnesium sulphate and submitted to the analysis by gas chromatography.

Analysis in gas chromatography

Samples for gas chromatography were prepared as described above and conversions of substrate into the corresponding products were determined by achiral GC based on relative peak areas. Ratios of regioisomers **31a/b**, **33a/b** and **34a/b** were calculated based on peak areas derived from chiral GC. Calculation of the ratios of **32a**/**b** and **35a/b** regioisomers was based on peak areas derived from achiral GC. Enantiomeric excesses (ee) of azido alcohol products **31**, **32** and **35** were determined using chiral GC and according to literature.[174]
Non-commercially available products were detected and confirmed using gas chromatography coupled with mass spectrometry (GC-MS) and NMR (see below). A summary of the detailed information of the different temperature programs and columns can be found in Table 13-15.

Table 13. GC programs used for the analysis of the listed compounds using the columns listed carried out at TU Braunschweig.

Compound	Temperature profile	RT (min)	Column
trans-2-hexene (**1**)	50 °C // 10 °C/min // 70 °C // 50 °C/min // 300 °C	3.6	Achiral: Optima 5s
trans-2,3-epoxyhexane (**21**)		8.4	
3-azido-2-hydroxyhexane (**31b**)		11.8	
2-azido-3-hydroxyhexane (**31a**)		11.8	
trans-2-heptene (**2**)	70 °C, 2 min // 50 °C/min // 125 °C, 3 min // 50 °C/min // 300 °C	2.1	
trans-2,3-epoxyheptane (**22**)		3	
trans-4,5-epoxyhexanol (**23**)		3.9	
trans-3,4-epoxyheptane (**24**)		5.6	
3-azido-2-hydroxyheptane (**32b**)		6	
2-azido-3-hydroxyheptane (**32a**)		6.1	
3,4-azidoheptanol (**34a** and **34b**)		6.3	
azidohexanediol (**33a** and **33b**)		7.3	
cyclohexene	100°C, 3min// 50°C/min // 200°C// 20°C/min// 300°C	2.4	
cyclohexene oxide		3.3	
trans-1-phenylpropylene oxide (**25**)	80 °C, 1 min // 10 °C/min // 160 °C // 20 °C/min // 200 °C	6.8	
1-azido-1-phenylpropan-2-ol (**35b**)		10.3	
2-azido-1-phenylpropan-1-ol (**35a**)		10.4	
trans-(2S,3S)-epoxyhexane (**21**)	30 °C // 2 °C/min// 50 °C // 10 °C/min// 200 °C	8	Chiral: HYDRODEX γ-DiMOM
trans-(2R,3R)-epoxyhexane (**21**)		8.3	
trans-(2S,3S)-epoxyheptane (**22**)	40 °C // 1 °C/min// 50 °C // 10 °C/min // 200 °C	12.1	
trans-(2R,3R)-epoxyheptane (**22**)		12.2	
(*2S,3R*)-2-azido-3-hydroxyhexane (**31a**)	70 °C, 20 min // 10 °C/min // 200 °C	25.6	Chiral: Lipodex E
(*2R,3S)*-2-azido-3-hydroxyhexane(**31a**)		25.8	
(*3R,2S*)-3-azido-2-hydroxyhexane (**31b**)		26.3	
(*3S,2R*)-3-azido-2-hydroxyhexane (**31b**)		26.6	
(*2S,3R*)-2-azido-3-hydroxyheptane (**32a**)	85 °C // 0.5 °C/min // 95 °C // 10 °C/min // 200 °C	17.5	
(*2R,3S)*-2-azido-3-hydroxyheptane(**32a**)		17.9	
(*3R,2S*)-3-azido-2-hydroxyheptane (**32b**)		19.6	
(*3S,2R*)-3-azido-2-hydroxyheptane (**32b**)		20.7	
4-azidohexane-1,5-diol (**34a**)	85 °C, 10 min // 5 °C/min // 170 °C, 5 min // 10 °C/min // 200 °C	24.3	
4-azidohexane-1,5-diol (**34a**)		24.6	
5-azidohexane-1,4-diol (**34b**)		25.1	
5-azidohexane-1,4--diol (**34b**)		25.2	

3-azido-4-hydroxyheptane (**33a**)	80 °C, 10 min // 2 °C/min // 110 °C, 5 min // 10 °C/min // 200 °C	22.1
4-azido-3-hydroxyheptane (**33b**)		22.1
4-azido-3-hydroxyheptane (**33b**)		22.8
(1*S*)-azido-1-phenylpropan-(2*R*)-ol (**35b**)	110 °C, 10 min // 10 °C/min // 165 °C, 10 min // 5 °C/min // 170 °C, 10 min // 5 °C/min // 205 °C	15.6
(1*R*)-azido-1-phenylpropan-(2*S*)-ol (**35b**)		15.8
(2*R*)-azido-1-phenylpropan-(1*S*)-ol (**35a**)		17.08
(2*S*)-azido-1-phenylpropan-(1*R*)-ol (**35a**)		17.15

Table 14. GC programs used for the analysis of the listed compounds using the columns listed carried out at TU Delft.

Compound	Temperature profile	RT (min)	Column
trans-2-hexene (**1**)	50 °C, 6 min // 30 °C/min // 115 °C, 1 min // 30 °C/min // 155 °C, 1min // 30 °C/min // 325 °C	3.4	Achiral:CP-sil 5
trans-2,3-epoxyhexane (**21**)		5.9	
2-amino-3-hydroxyhexane (**41a**)		8.3	
2-azido-3-hydroxyhexane (**31a**)		9.6	
trans-2-heptene (**1**)	60 °C, 5 min // 30 °C/min // 140 °C, 2 min// 30 °C/min // 325 °C	2.9	
trans-2,3-epoxyheptane (**22**)		3.2	
2-amino-3-hydroxyheptane (**42a**)		7.8	
2-azido-3-hydroxyheptane (**32a**)		8.8	

Table 15. GC programs used for the analysis of the listed compounds using the columns listed carried out at Enzymicals.

Compound	Temperature profile	RT (min)	Column
trans-2-hexene (**1**)	40 °C, 3 min // 20 °C/min// 90 °C // 20 °C /min // 180 °C, 2 min	3	Achiral: BPX5
trans-2,3-epoxyhexane (**21**)		4.7	
2-amino-3-hydroxyhexane (**41a**)		9.7	
3-amino-2-hydroxyhexane (**41b**)		9.9	
3-azido-2-hydroxyhexane (**31b**)		11.3	
2-azido-3-hydroxyhexane (**31a**)		11.3	
trans-2-heptene (**2**)	70 °C, 3 min // 15 °C/min // 135 °C, 3 min // 20 °C/min // 180 °C, 2 min	2.5	
trans-2,3-epoxyheptane (**22**)		4.9	
2-amino-3-hydroxyheptane (**42a**)		7.4	
3-amino-2-hydroxyheptane (**42b**)		7.5	
3-azido-2-hydroxyheptane (**32b**)		10.1	
2-azido-3-hydroxyheptane (**32a**)		10.4	
trans-(2S,3S)-epoxyhexane (**21**)	40 °C, 12 min // 20 °C min^{-1} // 180 °C	11.3	Chiral: HYDRODEX β-3p
trans-(2R,3R)-epoxyhexane (**21**)		11.9	
trans-(2S,3S)-epoxyheptane (**22**)	50 °C, 5 min // 2.5 °C min^{-1} // 60 °C, 10 min // 20 °C min^{-1} // 180 °C	15.2	
trans-(2R,3R)-epoxyheptane (**22**)		15.4	
(*2S,3R*)-2-azido-3-hydroxyhexane (**31a**)	70 °C, 20 min // 10 °C min^{-1} // 200 °C	26.6	Chiral: Lipodex G
(*2R,3S)*-2-azido-3-hydroxyhexane(**31a**)		26.7	
(*3R,2S*)-3-azido-2-hydroxyhexane (**31b**)		27.1	

(*3S,2R*)-3-azido-2-hydroxyhexane (**31b**)		27.1
(*2S,3R*)-2-azido-3-hydroxyheptane (**32a**)	85 °C // 0.5 °C min^{-1} // 95 °C // 10 °C min^{-1} // 200 °C	28.3
(*2R,3S)*-2-azido-3-hydroxyheptane(**32a**)		28.5
(*3R,2S*)-3-azido-2-hydroxyheptane (**32b**)		29.07
(*3S,2R*)-3-azido-2-hydroxyheptane (**32b**)		29.1

GC-MS analysis

GC-MS analysis of the compounds was performed on a Shimadzu GCMS-QP2010SE equipped with a ZB-5MS GUARDIAN column (Phenomenex) or BPX5 and helium as carrier gas.

TLC stains

Epoxides were stained with Ceric Phosphomolybdic acid stain. The staining solution was prepared by mixing 2g $Ce(SO_4)_2$ and 5g Phosphomolybdic acid ($H_3PMo_{12}O_{40}$) in 200 ml ddH_2O and 16 ml cc. H_2SO_4 which was added at last. The mixture was slightly heated up to obtain a clear yellow solution. For visualization, the dry TLC plate was dipped into the staining solution and heated using a heat pistol until the characteristic blue spots became visible.

Visualization of azido alcohols and alkenes (not volatile) was carried out using a potassium permanganate stain. The stain was prepared by mixing 2.4 g of $KMnO_4$ with 16 g K_2CO_3 in 240 ml ddH_2O, at last 4 mL of 5% (w/v) NaOH solution was added and the solution was mixed until all the potassium permanganate was dissolved. Staining and visualization were carried out as described above and yellow-brownish spots became visible on the pink background.

Ninhydrin stain was used for staining amino alcohols. 1 g of ninhydrin were mixed in 100 mL ethanol and 3 mL acetic acid were added at last. After heating the stained TLC plate purple spots became visible.

NMR analysis

NMR spectra were recorded either on a Bruker Avance II 300 MHz or on a Bruker Avance 600 MHz spectrometers equipped with an inverse $^{1}H/^{13}C/^{15}N/^{31}P$ quadruple resonance cryoprobe head and z field gradients. The sample was dissolved in deuterated chloroform ($CDCl_3$) and 1D proton, DEPT, COSY double quantum filter, and HSQC experiments were performed at 25°C. The regioisomers of chemically synthesised azido alcohols **31**, **33** and **34** were identified by NMR spectroscopy employing COSY and HSQC techniques and further compared to the assignment of regioisomers by GC and GC-MS.

Analytical data of the compounds

trans-2,3-epoxyhexane (**21**):

^{1}H NMR (300 MHz, $CDCl_3$) δ = 2.74 (qd, *J*=5.2, 2.3, 1H), 2.63 (td, *J*=5.4, 2.3, 1H), 1.54 – 1.40 (m, 1H), 1.29 (d, *J*=5.2, 1H), 0.95 (t, *J*=7.1, 1H).

trans-2,3-epoxyheptane (**22**):
^{1}H NMR (300 MHz, $CDCl_3$) δ = 2.74 (qd, *J*=5.2, 2.3, 1H), 2.62 (td, *J*=5.5, 2.3, 1H), 1.55 – 1.32 (m, 2H), 1.29 (d, *J*=5.2, 1H), 0.91 (t, *J*=7.1, 1H).

trans-4,5-epoxyhexanol (**23**):
^{1}H NMR (300 MHz, $CDCl_3$) δ = 3.90 (ddd, *J*=12.5, 5.6, 2.5, 1H), 3.67 – 3.54 (m, 1H), 3.02 – 2.83 (m, 2H), 1.98 (dd, *J*=6.9, 5.9, 1H), 1.62 – 1.35 (m, 4H), 0.95 (t, *J*=7.2, 3H).
trans-3,4-epoxyheptane (**24**):
^{1}H NMR (300 MHz, $CDCl_3$) δ = 2.72 – 2.61 (m, 2H), 1.61 – 1.40 (m, 6H), 0.99 (t, *J*=7.5, 3H), 0.96 (t, *J*=7.2, 3H).
trans-1-phenylpropylene oxide (**25**):
^{1}H NMR (300 MHz, $CDCl_3$) δ = 7.40 – 7.20 (m, 5H), 3.57 (d, *J*=2.0, 1H), 3.04 (dq, *J*=5.1, 1H), 1.45 (d, *J*=5.1, 3H).

Diastereomeric mixture of 2-azido-3-hydroxyhexane (**31a**) and 3-azido-2-hydroxyhexane (**31b**):
^{1}H NMR (600 MHz, $CDCl_3$) δ = 3.90 – 3.81 (m, 1H, **31b**), 3.65 – 3.58 (m, 1H, **31a**), 3.52 (dq, *J*=6.7, 3.9, 1H, **31a**), 3.42 – 3.35 (m, 1H, **31b**), 1.61 – 1.29 (m, 14H, **31a** and **31b**), 1.25 (d, *J*=6.7, 3H, **31a**), 1.19 (d, *J*=6.4, 3H, **31b**), 0.96 (t, *J*=7.2, 3H, **31b**), 0.94 (t, *J*=7.3, 3H, **31a**).
^{13}C NMR (151 MHz, $CDCl_3$) δ 73.55 (**31a**), 69.99 (**31b**), 67.92 (**31b**), 61.77 (**31a**), 34.57 (**31a**), 32.07 (**31b**), 19.65 (**31b**), 18.96 (**31a**), 18.13 (**31b**), 13.91 (**31a**), 13.80 (**31b**), 13.11 (**31a**).
ESI-HRMS: [M+Na$^+$] = 166.09524 m/z (calculated [M+Na+] = 166.09508 m/z).

2-azido-3-hydroxyheptane (**32a**)
^{1}H NMR (300 MHz, $CDCl_3$) δ = 3.66 – 3.47 (m, 2H), 1.51 – 1.28 (m, 6H), 1.26 (d, *J*=6.6, 3H), 0.92 (t, *J*=7.0, 3H).

Diastereomeric mixture of 5-azidohexan-1,4-diol (**33a**) and 4-azidohexan-1,5-diol (**33b**):
^{1}H NMR (600 MHz, $CDCl_3$) δ = 3.92 – 3.85 (m, 2H, **33b**), 3.82 – 3.60 (m, **33a** 3H, **33b** 1H), 3.50 – 3.44 (m, 1H, **33a**), 3.42 – 3.40 (m, J=4.98, 1H, **33b**), 2.86 (s, 1H, **33a**), 2.63 (s, 1H, **33b**), 2.34 (s, 1H, **33a**), 2.17 (s, 1H, **33b**), 1.65 – 1.34 (m, 8H, **33a** and **33b**), 0.97 (t, J=7.0, 3H, **33a**), 0.96 (t, J=6.9, 3H, **33b**).
^{13}C NMR (151 MHz, $CDCl_3$) δ 73.80 (**33a**), 72.39 (**33b**), 67.04 (**33b**), 64.64 (**33a**), 63.33 (**33a**), 62.60 (**33b**), 36.00 (**33b**), 32.83 (**33a**), 19.74 (**33a**), 19.01 (**33b**), 14.13 (**33b**), 14.03 (**33a**).

ESI-HRMS: $[M+Na^+]$ = 182.09010 m/z (calculated [M+Na+] = 182.0905 m/z).

Diastereomeric mixture of 3-azido-4-hydroxyheptane (**34a**) and 4-azido-3-hydroxyheptane (**34b**):
^{1}H NMR (600 MHz, $CDCl_3$) δ 3.72 – 3.64 (m, 1H, **34a**), 3.62 – 3.55 (m, 1H, **34b**), 3.42 – 3.32 (m, 1H, **34b**), 3.31 – 3.23 (m, 1H, **34a**), 1.73 – 1.31 (m, 14H, **34a** and **34b**), 1.04 (t, *J* = 7.4 Hz, 3H, **34a**), 1.00 (t, *J* = 7.4 Hz, 3H, **34b**), 0.96 (t, *J* = 7.1 Hz, 3H, **34b**), 0.95 (t, *J* = 7.2 Hz, 3H, **34a**).
^{13}C NMR (151 MHz, $CDCl_3$) δ = 75.64 (**34b**), 73.64 (**34a**), 69.46 (**34a**), 67.26 (**34b**), 34.70 (**34a**), 31.74 (**34b**), 25.58 (**34b**), 22.90 (**34a**), 19.93 (**34b**), 19.20 (**34a**), 14.20 (**34a**), 14.10 (**34b**), 11.24 (**34a**), 10.43 (**34b**).
ESI-HRMS: $[M+Na^+]$ = 180.11081 m/z (calculated [M+Na+] = 180.1112 m/z).

Obtained NMR data for **10**[175,176] were consistent with literature data.

2-amino-3-hydroxyhexane (**41a**):
^{1}H NMR (600 MHz, $CDCl_3$) δ = 3.69 – 3.60 (m, 1H), 3.55 – 3.37 2.74 (m, 3H), 3,15 – 3.06 (m, 1H), 1.60 – 1.27 (m, 4H), 1.10 (d, *J*=6.5, 3H), 0.94 (t, *J*=7.02, 3H).
^{13}C NMR (151 MHz, $CDCl_3$) δ = 73.04, 50.85, 34.73, 19.31, 15.23, 14.11.

2-amino-3-hydroxyheptane (**42a**):
^{1}H NMR (600 MHz, $CDCl_3$) δ = 3.50 – 3.39 (m, 1H), 3.10 – 2.86 (m, 1H), 1.95 – 1.70 (m, 3H), 1.55 – 1.18 (m, 6H), 1.01 (d, J=6.39, 3H), 0.91 (t, J=7.02, 3H).
^{13}C NMR (151 MHz, $CDCl_3$) δ = 76.63, 50.34, 32.14, 28.38, 22.82, 16.77, 14.06.

2.2.4 Biochemical and chemical reactions

Enzymatic epoxidations

Bioconversion using cytochrome P450 monooxygenases

Purified CYP154C5 or CYP154H1 were dissolved in 50 mM potassium phosphate buffer pH 7.5 to a final concentration of 3 µM (protein concentration determined by CO-difference spectroscopy[169]). PdR and Pdx concentrations were fixed to 3 and 16 µM respectively to ensure an electron transfer component activity of approximately 1 U/mL (determined by cytochrome c assay[170]). The substrate was added from a 500 mM stock solution in ethanol to a final concentration of 5 mM and the cofactor NADH (50 µM) was added. Formate dehydrogenase from *Candida boidinii* (0.5 U/mL) and sodium formate (150 mM) were applied for cofactor regeneration. The reaction was carried out in a final volume of 1.2 mL at 30 °C with magnetic stirring overnight. Finally, a 400 µL sample was prepared as described in section 2.2.3.

When whole cells were used cell pellets of *E. coli* C43 (DE3) overexpressing Pdx, PdR and CYP154C5 or CYP154H1 were resuspended in 50 mM potassium phosphate buffer pH 7.5. 20 mM glucose was added for cofactor regeneration and the reaction was carried out in final volume of 1.2 mL at 30 °C with magnetic stirring overnight. Control reactions were performed using the same biocatalyst with overexpressed Pdx and PdR only.

Chemical epoxidation catalysed by MsAct

The enzyme catalyses the perhydrolysis of a given ester to generate the relative peracid and alcohol. The peracid can then perform a spontaneous chemical epoxidation if a suitable substrate is present. For this 100 µg of purified MsAct were mixed in 200 mM potassium phosphate buffer pH 7.6 with different concentrations of ethyl acetate as ester (10, 20 and 30 % v/v) to a final volume of 1 mL. Cyclohexene from a 500 mM stock in ethanol was added to a final concentration of 10 mM and urea hydrogen peroxide (UHP) was supplemented in different proportions (1, 2, 3, 4, 10 and 2.5 x 4 equivalents) to slowly release hydrogen peroxide for the perhydrolysis. After 24 h incubation at 18 °C shaking at 900 rpm, 400 µL were taken and the sample was prepared for gas chromatography as described in the section 2.2.3.

Epoxidation catalysed by the unspecific peroxygenase from *Agrocybe aegerita*

The cell-free extract harbouring the unspecific peroxygenase from *A. aegerita* (42 µM) was kindly provided by Dr. Frank Hollman (TU Delft) and the standard protocol was adapted and applied to test *trans*-2-heptene (**2**) epoxidation.[63] The enzyme concentration was adjusted to 100 nM in 50 mM potassium phosphate buffer pH 7.5, **2** was added from a 500 mM stock in ethanol to a final concentration of 20 mM. Finally, H_2O_2 was added dropwise over 8 h to a final concentration of 100 mM (5 eq.). After 24 h at 30 °C and at 900 rpm shaking, an aliquot was extracted according to the sample preparation protocol described previously.

Styrene monooxygenase

General whole cells epoxidation

Small-scale biotransformations (1 mL) were performed in 50 mM potassium phosphate buffer at pH 7.5 containing 60 g/L (wet cell weight) of *E.coli* BL21 harbouring StyAB, 5 mM substrate (**1-2**) and 40 mM glucose for cofactor regeneration at 30 °C. When no oxygen was used, the reactions were carried out in a 4 mL glass reaction tube with magnetic stirring to provide enough oxygen present in atmospheric air. Using an oxygen balloon, reactions were carried out in 1.5 mL tube in a thermomixer. The reaction was stopped after 1h by incubating a 400 µL sample on ice and by immediately adding 400

µL of TBME containing 0.1% (v/v) dodecane as an internal standard. The sample was prepared and analysed by GC as described above.
Negative controls were carried out by using *E.coli* BL21 with an empty vector.
The optimal cellular concentration was determined by adjusting the same cell pellet to different concentrations 7.5, 15, 30, 60 and 120 g/L and the activity was determined as described in section 2.2.2 (enzyme assays).

General cell-free extract epoxidation

The cell-free extract containing StyAB concentration was adjusted to 9.2 g/L (total protein content) in 50 mM potassium phosphate buffer at pH 7.5. 5 mM substrate (**1** or **2**) was added and the commercial FDH from *C. boidinii* (0.5 U/mL) was added with 150 mM NaCOOH for cofactor regeneration. The reactions with or without oxygen were performed as described in the previous section and the samples for GC analysis were prepared as described before.

General cell-free extract epoxidation using BNAH for cofactor regeneration

Cell-free extracts of *E. coli* BL21 harbouring recombinant StyA were dissolved in 50 mM potassium phosphate buffer at pH 7.5 to a final concentration of 9.2 g/L (total protein content), 5 or 10 mM of either substrate **1** or **2** were added and 2 or 4 equivalents of BNAH were added for cofactor regeneration. $FADH_2$ was tested at different concentration (50 or 200 µM) to determine the optimal ratio. The reaction was magnetically stirred or shacked at 30°C for 1 h and samples taken for GC analysis were treated as described above.
StyA stability was measured leaving concentrated cell-free extracts at 30 °C with magnetic stirring for as long as 24 h. At the desired time points the appropriate volume was taken and transferred to a new reaction tube containing 10 mM BNAH, 50 µM $FADH_2$ and 3.5 mM *trans*-2-heptene (**2**) (from a 50 mM stock). The reaction was carried out for 5 min and the activity was measured according to the method described in section 2.2.2 (enzyme assays).

Design of Experiment (DOE)

The experimental design simplifies the optimization of complex systems by creating a statistical model of a given reaction. The model needs some inputs (variables' values) and outputs (responses). The aim is to determine which factors are significantly contributing to the response values. In this case, four different variables were considered (Table 16). The response factor considered was the concentration of epoxide **2** after 2 h. The design was done using the software RStudio with the DOE plug-in installed. Screening was the chosen option in the experimental designs and 8 runs with 4 factors and 2 replicates were selected for the design.

The reactions were performed in 20 mL 50 mM potassium phosphate buffer at pH 7.5 with 5 mM *trans*-2-heptene (**2**) with magnetic stirring. Glucose was used for cofactor regeneration at different concentrations, the rest of the settings used are shown in Table 16. After 2 h of reaction time, a 400 µL sample was taken and prepared for the analysis on gas chromatography.

Table 16 Factors and values used for low and high settings. *wcw = wet cell weight

Factor	Low setting	High setting
OD_{600}	10 (15 g/L wcw)	40 (60 g/L wcw)
Glucose	20 mM	100 mM
Stirring speed	150 rpm	400 rpm
Temperature	25 °C	30 °C

General epoxidation on 40 mL scale in 50/100 mL round bottom flask

To a solution of 50 mM potassium phosphate buffer at pH 7.5 and 5 mM substrate (**2**) (final concentration) 15 g/L (wcw) whole cells (*E. coli* BL21) harbouring StyAB were added. 2 or 4 equivalents of glucose were added and the reaction was magnetically stirred at 25 °C for 24 h. When applied, oxygen was provided through a double-necked round bottom flask where the O_2 filled balloon was fitted via a syringe fitting the diameter of the flask neck and sealed with PTFE tape. Samples for GC analysis were prepared as described before.

General epoxidation on 80 mL scale in 100 mL Schott bottle

Epoxidations using CFE of *E. coli* BL21 harbouring StyAB and LsADH (9.2 g/L, total protein content) were carried out in 50 mM potassium phosphate buffer at pH 7.5, *trans*-2-hexene (**1**) or *trans*-2-heptene (**2**) was added to a concentration of 50 mM from a 2 M stock in isopropanol (used as co-substrate for LsADH). CFE containing LkADH was added to a final concentration of 2.6 g/L (total protein content). Compressed air was provided through an HPLC frit with a 3 L/h flow. The reaction was mechanically stirred with overhead stirring for 24 h at 25 °C. Samples were periodically taken and analysed as described before.

When substrate feed was used, the appropriate amount of the 2 M stock was supplemented 5 times every hour until the final 50 mM concentration was reached.

When LkADH or LsADH were coexpressed only the cell-free extract harbouring StyAB and the ADH was used to a final concentration of 9.2 g/L (total protein content).

When enzyme feed was used, the appropriate amount from the stock CFE solution (75 g/L, total protein content) was supplement every hour for the first 5 hours until the final concentration of 9.2 g/L (total protein content) was reached. When a second organic phase was used (25% v/v) 400 µL samples were taken directly from the organic phase

containing 0.1% v/v dodecane. Afterwards, the samples were dried over anhydrous Na_2SO_4 and submitted to GC analysis.

Figure 6 shows the setup where reactions were run. A three-port lid was used to ensure correct spacing between overhead stirring and compressed air input. The third port can be used for enzyme and substrate feed, for sampling and as exhaust. The 100 mL Schott flasks are inside a water bath to maintain the temperature and the air was finely dispersed in solution via a porous metal frit used as sparger. The mixer attached to the overhead stirrer is positioned slightly higher than the frit to maximize the resident time of oxygen in solution.

Figure 6. The setup is composed of 100 mL Schott bottle with a three-port lid where overhead stirring and compressed air input can be inserted. The third port can be used for enzyme and substrate feed and for sampling.

Oxygen measurements and kinetics

The oxygen concentration was followed using an optic fibre oxygen sensor from Presens, connected to the respective meter (Microx 4, PreSens Precision Sensing GmbH). The meter shows the oxygen concentration in different units (µmol/L was chosen) and it was calibrated at 25 °C prior to each measurement to ensure accurate measurements.
Enzyme kinetics towards O_2 were determined according to a published protocol.[177] The enzyme rate derived from equation 1 (r_0). While the oxygen volumetric mass transfer coefficient without aeration present ($k_La'_0$) was estimated before starting the kinetics experiment using equation 2.

$$\frac{dC_{O_2}}{dt} = k_L a'_0 \cdot \left(C^*_{O_2} - C_{O_2}\right) - r_0 \qquad (1)$$

The overall oxygen volumetric mass transfer coefficient was estimated using the same reaction set up (4 mL glass reaction tube). The air was removed from 1 mL of 50 mM potassium phosphate buffer pH 7.5 containing CFE (2.3 g/L total protein content) using an N_2 flow until oxygen concentration dropped to 0. Afterwards, the nitrogen flow was stopped and the atmospheric oxygen was let flow back in the solution. The O_2 concentration curve over time was used to calculate the overall oxygen volumetric mass transfer coefficient (Eq. 2) using non-linear regression analysis.

$$\frac{dC_{O_2}}{dt} = k_L a' \cdot \left(C^*_{O_2} - C_{O_2}\right) \qquad (2)$$

When the overall oxygen volumetric mass transfer coefficient is estimated, the dependence of StyAB on the oxygen concentration can be determined using equation 1. The desired oxygen concentration was adjusted by using nitrogen to remove the proper amount of air. When the predetermined O_2 concentration was reached the enzyme solution was added and the rate at which the oxygen diffuses back in solution was measured. Since the oxygen is consumed by the enzyme activity as well it is possible to measure its initial rate but subtracting eq. 2 from eq.1. The first part of the curve (linear) was used to estimate the enzyme activity. It was appreciable that at lower oxygen concentrations the diffusion rate was faster than the enzyme rate.

Epoxidation on 200 mL scale in 250 mL Schott bottle (100 mM substrate)

The last scale up set up was similar to the previous protocol where the cell-free extract harbouring StyAB and LsADH was hourly fed to a final concentration of 9.2 g/L. 2 M stock in isopropanol of alkene **2** was fed hourly to a final concentration of 100 mM. 50 mM potassium phosphate buffer pH 7.5 was used as buffer and heptane (25% v/v) was used as second phase. The reaction was mechanically stirred at 300 rpm in a water bath at 25 °C for 24 h. Samples were taken directly from the organic phase containing 0.1% v/v dodecane, dried over anhydrous Na_2SO_4 and submitted to GC analysis.

Chemical epoxidations

Jacobsen epoxidation

The Jacobsen's catalyst is generally designed for terminal alkenes but as it was readily available it was tested using the standard protocol.[40] To 10 mM of *trans*-2-heptene (**2**) in saturated $NaHCO_3$ pH 8.0 at room temperature, 0.2 equivalents of either (*R,R*)- or (*S,S*)-Jacobsen's catalyst were added. 2 equivalents of NaOCl were added in portions to support the epoxidation. After 24 h incubation at 900 rpm shaking, a sample of 400 µL was taken and prepared and analysed by GC as described before.

Shi epoxidation

The typical Shi epoxidation was done in either 2 or 50 mL scale using different acetonitrile concentrations (5, 10 and 20% v/v). The test reaction was carried out in ddH_2O to a final volume of 2 mL. 2 mM of the Shi epoxidation catalyst were used to convert 10 mM of alkene **2** using 2 equivalents of Oxone© added in portions over the first 2h to start the catalytic cycle at room temperature for 24 h at 900 rpm shaking, $NaHCO_3$ was added to maintain the pH 8.0. The optimization was carried out in 50 mL scale using K_2CO_3 to maintain the pH at 11. Different concentrations of alkene **1** or **2** were used taking care of always using 0.2 equivalents of the Shi epoxidation diketal catalyst and 4 equivalents of Oxone© added in portions over the first 8 h together with K_2CO_3 to keep the pH constant. At fixed time points, a sample was taken and prepared as described in section sample preparation.

Epoxide ring opening by HHDHs and rational design of HHDHs

General biotransformation experiment

Small-scale biotransformations (1 mL) were performed in 50 mM potassium phosphate buffer at pH 7.5 containing 150 µg/mL of purified HHDH, 5 mM substrate (**21-25**) and 20 mM sodium azide (NaN_3) at 25 °C. After 15 h (substrates **21-24)** or 4 h (substrate **25**) incubation samples were taken for GC analysis. Chemical background azidolysis was monitored in reactions using the same reaction conditions but omitting HHDH enzyme. Alternatively, whole cells (60 g/L wet cell weight) or cell-free extract (9.2 g/L total protein content) of *E. coli* BL21 harbouring the desired HHDH were used to convert 10 mM or higher concentrations (50-300 mM) with respectively 40 mM or 2 equivalents of NaN_3 at 25 °C. Generally, after 4h and 24 h, a sample was taken and extracted as described above.

For conversions of cyclohexene oxide or 2-chlorocyclohexanol 20 mM of each substrate were used with 40 mM NaN_3 or NaCN as nucleophile. The catalyst containing HheG and the different mutants was used as whole cells (60 g/L wet cell weight) in 50 mM Tris·SO_4 pH 8 in a final volume of 2 mL. The reactions were incubated at 22 °C and mixed with magnetic stirring at 500 rpm. At different time points, 400 µL samples were taken and analysed by gas chromatography.

Preparative scale biotransformation

Preparative-scale conversions of **21** and **22** were performed in 40 mL of 50 mM potassium phosphate buffer at pH 7.5. Whole cells of *E. coli* BL21 harbouring either HheE or HheE5 (60 g/L wet cell weight) were used to convert 50 mM of either **21** or **22** and 100 mM sodium azide (NaN_3) were used as nucleophile. After 24 h at 25 °C the reaction mixture was extracted with 40 mL TBME and filtered through Celite®. The

extracted organic layer was washed with brine (1×) and MilliQ water (1×), dried over Na_2SO_4 and filtered before solvent evaporation. 142 mg (50% yield) of **31a** and 151 mg (48% yield) of **32a** were recovered after column chromatography using cyclohexane/diethyl ether, 90:10.

Docking studies

Molecular docking of MM2-force field energy minimized (*S*,*S*)-, (*R*,*R*)-**21**-**25** into the active site of a monomeric representation of HheC (PDB-ID: 1ZMT) HheD2 (unpublished) and HheG (PDB-ID: 5O30) was performed using the AutoDockVina algorithm present in YASARA Structure.[143,144] All ten substrates were docked into a simulation cell (X size = 25 Å, Y size = 25 Å, Z size = 25 Å; angles: α = 90°, β = 90°, γ =90°) around the following four residues: K91, L155, E197 and E216 for HheC the corresponding residues L96, Y160, K198 and E218 for HheD2 and the corresponding residues, D111, I175, E214 and E233 in HheG. For each substrate, 999 docking runs were performed while all atoms of the corresponding substrates were set as rigid with the only exception of substrate **3** and HheG where a flexible ligand was necessary to obtain one cluster with productive bind. The structure for each cluster was then energy minimized using YASARA2-force field after inserting the molecule of water as fixed element. This water molecule represents the nucleophile in its binding pocket and it is present in the respective crystal structures.

***In-silico* mutagenesis**

The “coupled moves” protocol allows the user to improve the enzyme-ligand interaction of a given structure or to radically change the substrate specificity by manually docking a non-natural substrate into the enzyme active site.[149] As there is no available crystal structure of HheG with bound substrate or nucleophile in the active site, the cyclohexene oxide, and (1*S*,2*R*)-2-cyanocyclohexanol were manually docked using available HheC crystal structures with styrene oxide or *(R)*-4-cyano-3-hydroxybutyrate in the active site. In detail, the respective ligands were aligned taking care to overlap the important moieties (i.e. epoxide ring, OH- and CN- groups) to the known crystal after superimposition of HheG over HheC co-crystallised with the substrate (1ZMT for styrene oxide and 4IXT for (1*S*,2*R*)-2-cyanocyclohexanol). The ligand rotamer library was generated by the free online tool Frog2[178] with default options, as these are rigid ligands and only one rotamer was generated. The same published setting[149] (command line) to run the coupled moves were used and are reported below:

~/Rosetta/main/source/bin/coupled_moves.linuxgccrelease
–s pdb_file –resfile res_file –database ~/Rosetta/main/database –mute
protocols.backrub.BackrubMover –extra_res_fa params_file –ex1 –ex2

–extrachi_cutoff 0 –nstruct 20 –coupled_moves::mc_kt 0.6 –coupled_moves::ntrials 1000 –coupled_moves::initial_repack false –coupled_moves::ligand_mode true –coupled_moves::ligand_weight N

The assisted mutagenesis produces typically 2000 unique sequences that were combined into a single FASTA file and submitted to the online tool WebLogo for visualization.[179]

Hydrogenation

Azides hydrogenation using heterogeneous palladium catalyst kit I

2-azido-3-hydroxyhexane (**31a**) or 2-azido-3-hydroxyheptane (**32a**) was dissolved in 4 mL of methanol containing 0.1% dodecane as internal standard to a final concentration of 5 mM, palladium on the different supports (1 mg/mL) was added and hydrogen was bubbled in the reaction mixture for 5 min prior to close the vial with the lid pierced by a needle with the H_2 balloon attached. After 30 and 60 min, an aliquot of 400 µL was taken, dried over $MgSO_4$ and directly submitted to GC analysis.

Preparation of Pd@apoferritin nanoparticles

Apoferritin from horse spleen (AFHS, Sigma Aldrich) was dissolved in EPPS buffer (100 mM EPPS, 10 mM NaCl, pH 8) to a concentration of 5 mg/mL. K_2PdCl_4 stock solution (50 mM in EPPS buffer) was added dropwise until the ratio with the 24er (number of AFHS monomers required for the assembly of the cage to store iron) was 250. The mixture was stirred overnight at room temperature to allow the Pd to diffuse inside the apoferritin complex. Afterwards, the palladium in excess was removed using a PD-10 column and the nanoparticles were formed by reducing Pd^{2+} using H_2. At first, hydrogen was bubbled through the yellowish solution for 15 min to remove all the oxygen present and subsequently the reaction was stirred for additional 2 h with the hydrogen balloon still connected but with the lid screwed to avoid bubbling. The nanoparticle formation can be appreciated as the liquid gradually turn dark grey testifying palladium aggregation. The nanoparticles were stored at 4 °C in a sealed tube.

The microscopic characterization was kindly carried out by Dr. Wuyuan Zhang (TU Delft) diluting the samples ten times with ddH_2O and using uranyl acetate as stain to allow visualization with TEM.

The hydrogenation was carried out by dissolving azide **32a** in 50 mM potassium phosphate buffer pH 7.5 to a final concentration of 10 mM and the proper amount of nanoparticles to have 3.5 µM of Pd was added. Hydrogen was bubbled for the first 5 min to remove all the oxygen present and subsequently the reaction was stirred for additional

48 h with the hydrogen balloon still connected but with the lid screwed. Samples were taken after 4, 24 and 48 h and treated as described previously prior to GC analysis.

Cascade reactions

Peak assignment of the diastereoisomers of the azido alcohols (31, 32, 35)

The assignment of the four peaks deriving from the unselective azidolysis of epoxides (**21**, **22** and 25) was achieved by conversion of enantiopure (*S,S*)-**21**, (*S*,*S*)-**22,** (*S*,*S*)-**25** and (*R*,*R*)-**25** using each a regioselective and a non-regioselective HHDH. To do this, HheG and HheD5 were used as non-stereoselective HHDH while HheE and HheE5 were used as regioselective enzymes. (*S,S*)-**21** and (*S*,*S*)-**22** were selectively produced in 50 mM potassium phosphate buffer at pH 7.5 by using CFE harbouring StyAB and FDH as cofactor regeneration (as described previously). After 24 h reaction, 150 µg/mL HHDH and 20 mM sodium azide at 25 °C were added. After 15 h, 400 µL sample was taken and prepared for GC analysis. By using non-regioselective enzymes (HheG for **21** and HheD5 for **22**), (2*R*,3*S*)-**31a**, (2*S*,3*R*)-**31b**, (2*R*,3*S*)-**32a** and (2*S*,3*R*)-**32b** were produced while only (2*R*,3*S*)-**31a** and (2*R*,3*S*)-**32a** were produced by using regioselective HHDH (HheE for **21** and HheE5 for **22**). GC-MS analysis of the same material was performed to assign the different regioisomers of **31** and **32**.

A similar strategy was used to assign the peaks of azido alcohols **35a** and **35b**. This time, (*S,S*)-**25** and (*R,R*)-**25** were commercially available. The azidolysis of (*S,S*)-**25** to an excess of (1*R,*2*S*)-1-azido-1-phenylpropan-2-ol (**35b**) next to a small amount of (1*S*,2*R*)-2-azido-1-phenylpropan-1-ol (**35a**) while ring-opening of the opposite enantiomer (*R,R*)-**25** generated (1*S,*2*R*)-1-azido-1-phenylpropan-2-ol (**35b**) in excess and a small amount of (1*R*,2*S*)-2-azido-1-phenylpropan-1-ol (**35a**) due to the S_N2-type mechanism and the observed regioselectivity of the chemical azidolysis.

Catalysts cross-interactions

Epoxide ring opening in presence of heptane

The influence of heptane (25% v/v) on the halohydrin dehalogenase was tested in 50 mM potassium phosphate buffer at pH 7.5. 10 mM substrate (**21** or **22**) and 50 mM sodium azide (NaN_3) were added and the solution was magnetically stirred at 25 °C, 600 rpm. After 4h and 15 h incubation, samples were taken for GC analysis. The biocatalyst was used as whole cells (60 g/L wet cell weight) harbouring HheE (substrate **21**) or HheE5 (substrate **22**).

Epoxide ring opening in presence of acetonitrile and Oxone©

The influence of acetonitrile (5%, 10% or 20 % v/v) and Oxone (50 mM or 10 mM) on the halohydrin dehalogenases was tested in 50 mM potassium phosphate buffer at pH 7.5. 10 mM substrate (**21** or **22** respectively for HheE and HheE5) and 40 mM sodium azide

(NaN_3) were added and the solution was magnetically stirred at 25 °C, 600 rpm. After 4 and 15 h incubation, samples were taken for GC analysis. The biocatalysts were used as whole cells (60 g/L wet cell weight) or cell-free extract (9.2 g/L total protein content) harbouring HheE (substrate **21**) or HheE5 (substrate **22**).

Epoxidation and ring opening in presence of Pd/C

The effects of Pd/C on the epoxidation carried out by the styrene monooxygenase was carried out by adding 1 mg/mL of heterogeneous catalyst to the epoxidation reactions carried out in 1 mL scale of either 50 mM potassium phosphate buffer pH 7.5. Whole cells (30 g/L wcw) or cell-free extract (4.6 g/L total protein content) harbouring respectively StyAB and StyA were used as catalyst to convert 5 mM of alkene **2** and respectively glucose and BNAH were use as cofactor regeneration. After 1 h, the reaction was sampled for GC analysis as described before.

1 mg/mL of Pd/C was also added to a solution where whole cells harbouring HheE5 (60 g/L wet cell weight) and 5 mM of epoxide **22** were present. 20 mM of NaN_3 was supplemented as nucleophile and the reactions were carried out for 1h at 25 °C in a thermomixer at 800 rpm. After 1 h, the reactions were sampled for GC analysis as described before.

Compatibility of the hydrogenation catalyst with the enzyme catalysts

The effect of the presence of cell-free extract, whole cells, BNAH and $FADH_2$ was studied as well. The hydrogenation was performed in 50 mM potassium phosphate buffer pH 7.5 where Pd/C (1 mg/mL) with 5 mM of racemic 2-azido-3-hydroxyheptane (**32a**) as substrate. In separate experiments Cell-free extract (4.6 g/L total protein content), whole cells (60 g/L wcw) BNAH (10 mM) and $FADH_2$ (200 µM) were added to assess to which extent the reaction is slowed down. After 15, 30 and 60 min an aliquot of 400 µL was taken and 10 µL of a solution of 250 mM KOH were added. The sample was extracted with the same volume of *tert*-butylmethylether (TBME) containing 0.1 % v/v dodecane as internal standard. The resulting organic phase was dried over anhydrous $MgSO_4$ prior to injection on GC.

Cascade reaction using StyAB, HHDH and Pd/C

The epoxidation of *trans*-2-heptene (**2**) (fed hourly to a final concentration of 20 mM) was carried out in 60 mL (+20 mL of heptane) of 50 mM potassium phosphate buffer pH 7.5. CFE harbouring StyAB and LsADH was added hourly to a final concentration of 9.2 g/L (total protein content). After 24 h, cell-free extract containing HheE5 and 50 mM of sodium azide was added in one portion to start the epoxide ring opening. After 48 h of total reaction time, Pd/C was added (2 mg/mL) and 3 equivalents of triethylsilane (TES) were supplemented in portions over 8 h to generate H_2 *in situ*. After 72 h of reaction, the

mixture was basified to pH 14 using NaOH pads and saturated with NaCl, the reaction mixture was divided into four falcon tubes and extracted 4 times with TBME. The phases were separated by centrifuging 2 min at 6000 rpm. The organic layers were combined, dried over anhydrous Na_2SO_4 and the solvent was removed in vacuum. The enantiopure (2*R,*3*S*)-2-amino-3-hydroxyheptane (**42a**) was purified by column chromatography as described previously yielding 3.8 mg of pure product (2.4% isolated yield).

Alternatively, giving the low yield of the full cascade the last step was separated. The first 48 h of the cascade were performed as described above using either HheE or HheE5 if alkene **1** or **2**, respectively were used as substrates in the first step. Afterwards, the reaction mixture was basified and saturated with NaCl. The mixture was divided into four falcon tubes and extracted 4 times with TBME. The phases were separated by centrifugation for 2 min at 6000 rpm. The organic layers were combined, dried over anhydrous Na_2SO_4 and the solvent was removed in vacuum. The crude extract was either submitted to column chromatography to purify enantiopure (2*R,*3*S*)-2-azido-3-hydroxyhexane (**41a**) or (2*R,*3*S*)-2-azido-3-hydroxyheptane (**42a**) or dissolved in 20 mL methanol for further reduction. The reduction was performed at RT and magnetically stirred at 500 rpm, hydrazine (100 mM) was added dropwise to prevent the generation of heat. After approx. 4 h the reaction mixture was filtered through Celite® and the solvent was removed in vacuum. The pure amino alcohols were obtained by column chromatography using a mixture of chloroform:methanol:ammonia (90:9:1).

Cascade reaction using Shi epoxidation diketal catalyst, HHDH and Pd/C

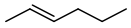
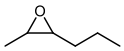
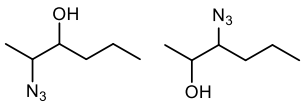
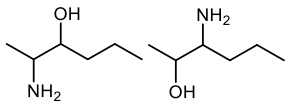
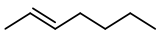
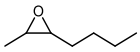
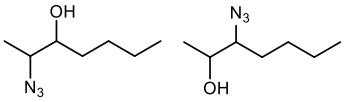
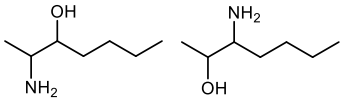
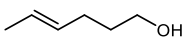
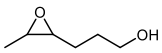
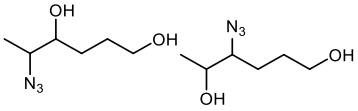
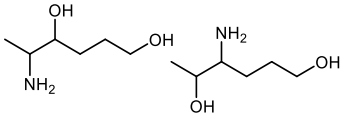
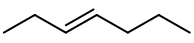
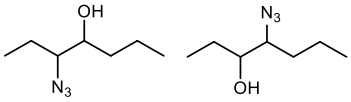
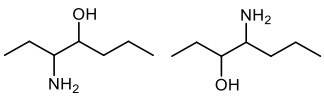
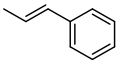
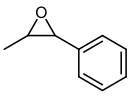
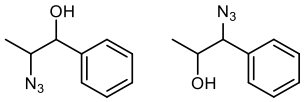
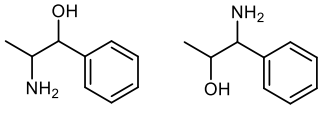
The first step was carried out at -2 °C in 47.5 mL ddH_2O with 2.5 mL of acetonitrile (5% v/v). 0.2 equivalents of Shi epoxidation were used to convert 50 mM of alkene **2.** 4 equivalents of Oxone© were added in portions over the first 8 h and solid K_2CO_3 was used to maintain the pH at 11 after addition of Oxone©. After incubation for 24 h at 300 rpm, the temperature was adjusted to 25 °C before lowering the pH to 7.5 using 95 % H_2SO_4. After the reaction conditions were adjusted, whole cells (60 g/L wcw) harbouring HheE5 were added together with 50 mM of NaN_3 and the reaction was continued for additional 24 h. After 48 h of total reaction time, Pd/C was added (2 mg/mL) and 3 equivalents of triethylsilane (TES) were supplemented in portions over 8 h to generate H_2 *in situ*. After 72 h of reaction, the mixture was basified to pH 14 using NaOH pads and saturated with NaCl, the reaction mixture was divided into four falcon tubes and extracted 4 times with TBME. The phases were separated by centrifuging 2 min at 6000 rpm. The organic layers were combined, dried over anhydrous Na_2SO_4 and the solvent was removed in vacuum. The enantiopure (2*S,*3*R*)-2-amino-3-hydroxyheptane (**42a**) was purified by column chromatography as described previously yielding 10.9 mg of pure product (3% isolated yield).

As for the cascade where StyAB is used in the first step, it was necessary to separate the hydrogenation giving the low yield of the complete cascade. The first 48 h of the cascade were performed as described above taking care of using the correct HHDH depending on the substrate used in the epoxidation. Afterwards, the reaction mixture was basified saturated with NaCl. The mixture was divided into four falcon tubes and extracted 4 times with TBME. The phases were separated by centrifugation for 2 min at 6000 rpm. The organic layers were combined, dried over anhydrous Na_2SO_4 and the solvent was removed in vacuum. The crude extract was either submitted to column chromatography to afford (2*S*,3*R*)-2-azido-3-hydroxyhexane (**41a**) or (2*S*,3*R*)-2-azido-3-hydroxyheptane (**42a**) with a fairly good ee% or dissolved in 20 mL methanol for further reduction. The reduction was performed at RT and magnetically stirred at 500 rpm, hydrazine (100 mM) was added dropwise to prevent the generation of heat. After approx.4 h the reaction mixture was filtered through Celite® and the solvent was removed in vacuum. The pure amino alcohols were obtained by column chromatography using a mixture of chloroform:methanol:ammonia (90:9:1).

3. Results

This chapter is divided into two main sections; the first section (section 3.1) presents the data concerning the setup and optimization of two cascades converting linear alkenes to the corresponding aminoalcohols. Section 3.2 deals with a third cascade starting from a cyclic alkene. All cascades are composed of three steps: epoxidation, epoxide ring opening using azide as nucleophile and hydrogenation of the azido group. All alkenes substrates and corresponding epoxides, azido alcohols and aminoalcohols used in this work are visible in Table 17.

Table 17. List of compounds used in this work.

Alkenes	Epoxides	Azido alcohols	Amino alcohols
 (**1**) *trans*-2-hexene	 (**21**) *trans*-2,3-epoxyhexane	 (**31a**) 2-azido-3-hydroxyhexane, (**31b**) 3-azido-2-hydroxyhexane	 (**41a**) 2-amino-3-hydroxyhexane, (**41b**) 3-amino-2-hydroxyhexane
 (**2**) *trans*-2-heptene	 (**22**) *trans*-2,3-epoxyheptane	 (**32a**) 2-azido-3-hydroxyheptane, (**32b**) 3-azido-2-hydroxyheptane	 (**42a**) 2-amino-3-hydroxyheptane, (**42b**) 3-amino-2-hydroxyheptane
 (**3**) *trans*-4-hexen-1-ol	 (**23**) *trans*-4,5-epoxyhexanol	 (**33a**) 5-azidohexane-1,4-diol, (**33b**) 4-azidohexane-1,5-diol	 (**43a**) 5-aminohexane-1,4-diol, (**43b**) 4-aminohexane-1,5-diol
 (**4**) *trans*-4-heptene	 (**24**) *trans*-3,4-epoxyheptane	 (**34a**) 3-azido-4-hydroxyheptane, (**34b**) 4-azido-3- hydroxyheptane	 (**44a**) 3-amino-4-hydroxyheptane, (**44b**) 4-amino-3-hydroxyheptane
 (**5**) *trans*-*β*-methyl styrene	 (**25**) *trans*-1-phenylpropylene oxide	 (**35a**) 2-azido-1-phenylpropan-1-ol, (**35b**) 1-azido-1-phenylpropan-2-ol	 (**45a**) 2-amino-1-phenylpropan-1-ol (**45b**) 1-amino-1-phenylpropan-2-ol

3.1 Cascades starting from acyclic alkenes

3.1.1 First step: Epoxidation of vicinally di-substituted alkenes

3.1.1.1 Cloning of styrene monooxygenase from *Rhodococcus sp.* ST-10

Previous characterization of the styrene monooxygenase from *Rhodococcus sp.* ST-10 demonstrated significant activity of this enzyme on the chosen substrates (**1** and **2**).[68] Therefore, codon-optimized synthetic genes coding for the two subunits, StyA and StyB, have been ordered from the company GeneArt (Thermo Fisher Scientific division) for expression in *Escherichia coli*. The pETDuet vector was selected as expression vector as it harbours two multiple cloning sites (MSC). Both MSCs are controlled by the *E. coli*–bacteriophage T7 RNA polymerase/promoter.[180] *StyA* and *styB* genes were ordered as a 5'-fusion construct (Figure 7) with a single NdeI restriction site in between. Moreover, the HindIII restriction site at the 3' end of *styA* was present two times in the pETDuet vector. Hence, the best cloning strategy (Figure 7) was to modify the restriction sites at both *styA* ends by PCR. The *styB* gene was first cloned into the MSC-2 using NdeI and XhoI as restriction enzymes. After transformation of chemically competent *E.coli* DH5α, the successful insertion was verified by colony PCR (Figure 8, C) and sequencing. At the same time, *styA* was amplified via PCR to change the NdeI restriction site at the 5' end to NcoI and the HindIII restriction site at the 3' end to BamHI. The *styA* gene and pETDuet with styB present in MSC-2 were both double digested with the above-mentioned restriction enzymes and ligated in two different cloning experiments. After transformation into chemically competent *E.coli* DH5α, successful insertion was verified by colony PCR (Figure 8, B) and sequencing. Additionally, digestion of the newly generated pETDuet-styAB using NdeI, XhoI, NcoI and BamHI as restriction enzymes was performed to ensure the presence of both genes (Figure 8, A).

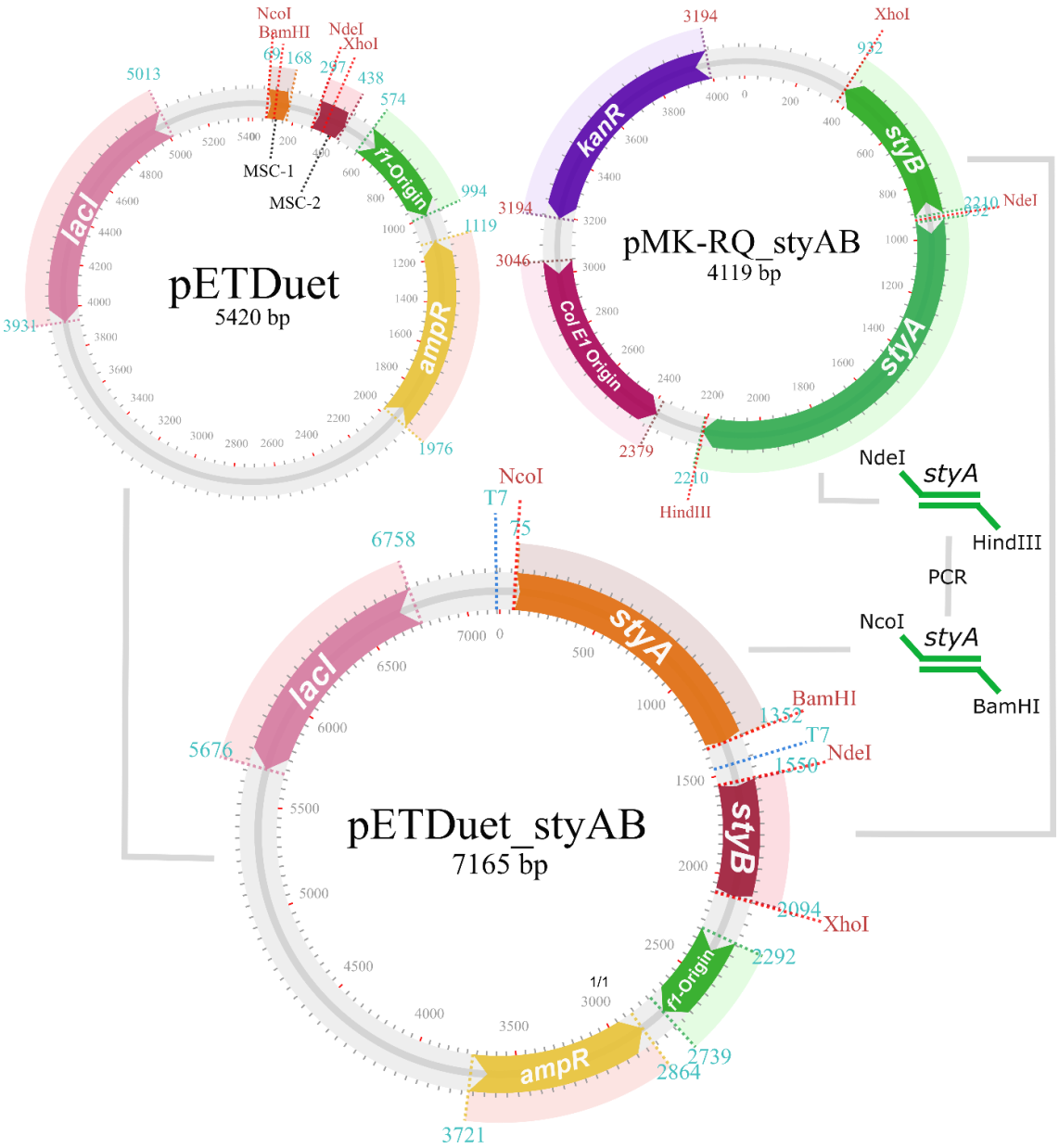


Figure 7. Cloning of *styA* and *styB* in MSC-1 and MSC-2, respectively, of the pETDuet expression vector. Restriction sites of *styA* were modified by PCR using primers with a mismatch to fit the corresponding endonucleases.

In Figure 8 (A), the result of the digestion using purified vectors from three different clones is presented. Digested vectors of clones 6 and 7 show visible bands at the expected sizes (1.2 kb for sytA and 0.5 kb for styB) and therefore, both were submitted to sequencing for validation. Additionally, styA was also sub-cloned into the pRSFDuet-1 with styB already present in its MSC-2. This time, *styA* was cut from pETDuet-styAB using NcoI and BamHI as restriction enzymes and inserted in pRSFDuet-styB cut with the same set of enzymes. This additional vector was created to co-express StyAB with the alcohol dehydrogenase from *Lactobacillus kefir* (LkADH) encoded on a pET21(+) vector. Moreover, the alcohol dehydrogenase from *Leifsonia sp.* S749 (LsADH) was also

tested since it had been reported as a regeneration system for StyAB.[68,85] Hence, the LsADH-encoding gene was ordered cloned in the pCDF-Duet-1 vector as it would be compatible with both pETDuet and pRSFDuet-1 vectors both containing *styA* and *styB*.

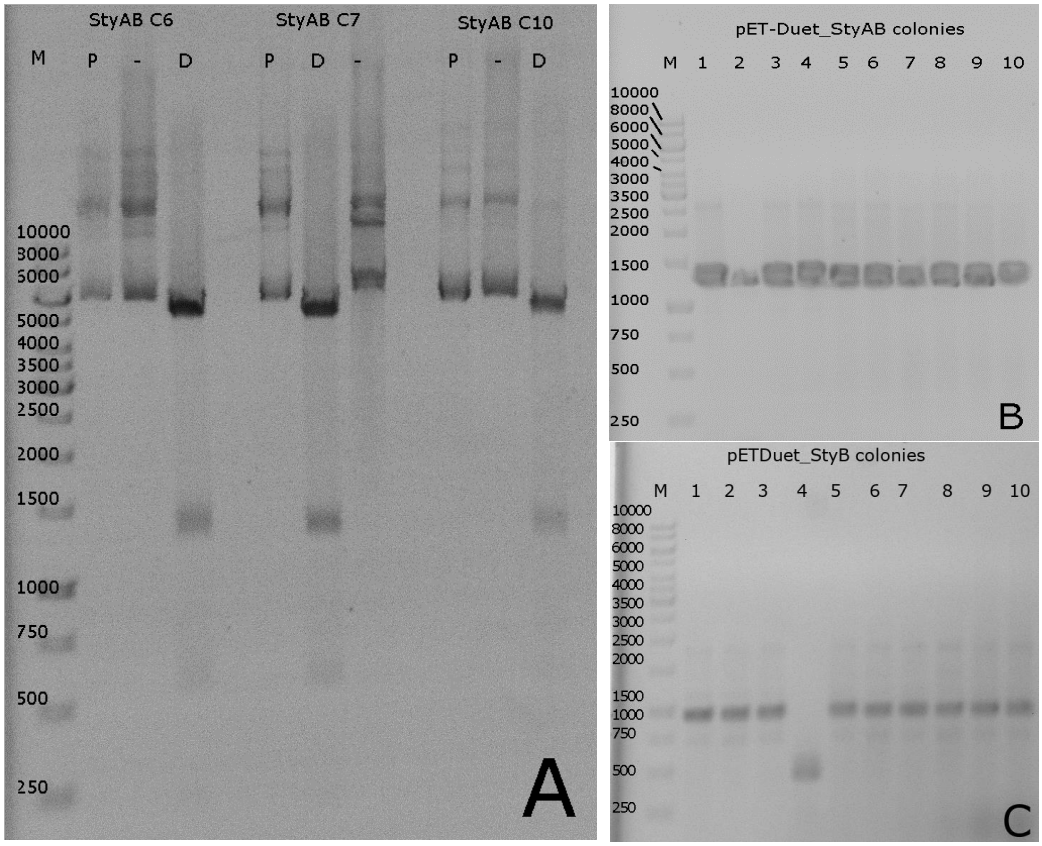


Figure 8. Gel A shows the digestion on three different colonies using NdeI, XhoI, NcoI and BamHI as restriction enzymes. Line "D" shows the digested samples and around 1.2 kb, *styA* is substantially more visible than the *styB's* band (0.5 kb). Lines "P" and "-" both represent the undigested plasmid although on the "-" samples the digestion buffer was also added to control for possible DNase contaminations. The successful insertion of *styA* in pETDuet-styB verified by colony PCR is shown in gel B. Gel C shows the colony PCR to confirm the insertion of *styB*. A stronger band at 1 kb and a faint band at 750 kb due to the forward primer that binds in the T7 promoter area of both MSC verify successful insertion.

3.1.1.2 Expression of styrene monooxygenase from *Rhodococcus sp.* ST-10

Different conditions for the heterologous expression of StyAB were tested to achieve best expression levels and simultaneously ensuring the highest activity. *E.coli* BL21 (DE3) and *E. coli* C43 (DE3) were used as expression hosts. Expression using BL21 (DE3) cells was carried out at 20 °C, whereas for C43 (DE3) 25 °C and 30 °C were tested. Additionally, both *Luria-Bertani* (LB) as well as *Terrific Broth* (TB) media were tested to analyse a possible influence on the expression. Lastly, the expression of StyAB in presence or absence of the chaperones was studied. To do this, the expression of

styrene monooxygenase was carried out with a second plasmid encoding the chaperone proteins (pG-KJE8).[74]

Two exemplary SDS-PAGE gels of StyAB expression are shown in Figure 9 together with the corresponding StyAB activities (Table 18). The activity was determined by measuring the epoxide formation over 5 min as described by Panke *et al.*,[78] in this case *trans*-2-heptene was used as substrate instead of styrene.

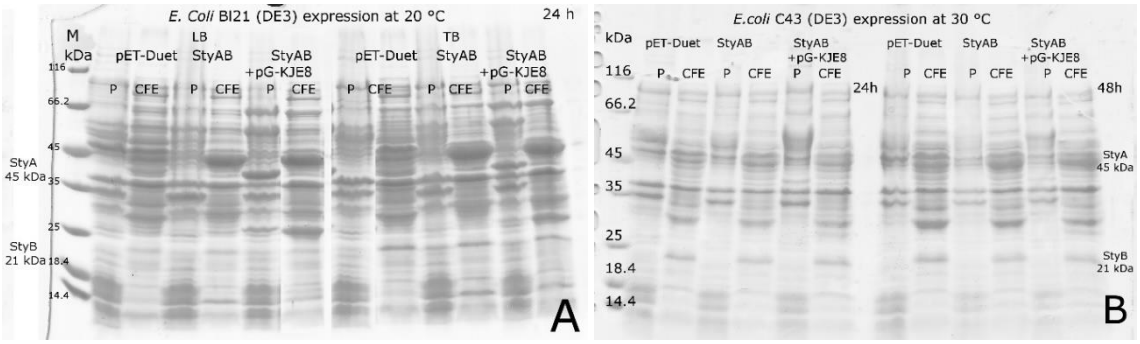


Figure 9. Exemplary gels showing the expression levels of StyAB in *E.coli* BL21 (DE3) (A) and *E.coli* C43 (DE3) (B). The table shows the activity measured for the different combinations of strain, temperature and IPTG concentration. The activity was measured determining the *trans*-2,3-epoxyheptane formation over 5 min.[78] All measurements were performed using whole cells with a concentration of 60 g/L (wet cell weight).

Overall, higher StyAB activity was obtained for expression in *E. coli* BL21 (DE3). Furthermore, co-expression of chaperons from vector pG-KJE8 had a positive effect on enzyme activity, perhaps allowing more StyB to be correctly folded during expression. However, the difference in activity between the samples with and without chaperones is not as remarkable as for the influence of the chosen expression host. Here, 2-fold higher activity was obtained with BL21 (DE3) as host. In contrast, increasing the IPTG concentration from 0.2 mM to 0.5 mM gave no or only a very limited improvement in activity. The best expression was achieved with *E. coli* BL21 (DE3) as host at 20 °C using 0.2 mM IPTG for induction. Co-expression of the chaperones did not appear crucial for StyAB expression even though slightly higher activities could be obtained.

Table 18. The best expression conditions of StyAB were determined using different expression host, temperature, IPTG concentration and presence of chaperon system

Strain	Catalyst	T (°C)	[IPTG] (mM)	mU	mU/mg (wet cell weight)

C43	StyAB,	25	0.5	59.7	0.995
C43	StyAB + pG-KJE8	25	0.5	75.52	1.26
C43	StyAB	30	0.5	79.18	1.32
C43	StyAB + pG-KJE8	30	0.5	82.84	1.38
BL21	StyAB	20	0.2	178.3	2.97
BL21	StyAB + pG-KJE8	20	0.2	180.2	3.00
BL21	StyAB	20	0.5	180.6	3.01
BL21	StyAB + pG-KJE8	20	0.5	168.6	2.81

(1 U = 1 µmol/min)

In addition, co-expression of StyAB with a cofactor regeneration system was tested to provide sufficient NADH for biocatalytic reactions. As mentioned previously, two different ADHs were ordered as synthetic genes to test the encoded proteins for cofactor regeneration. LkADH appears to have a higher overexpression compared to LsADH (Figure 10). On the other hand, more StyAB seems to be present when co-expressed together with LsADH from the pETDuet vector. Likely, the high overexpression of LkADH has a direct negative effect on the monooxygenase expression. In contrast, a more balanced expression was obtained for StyAB and LsADH co-expression, especially when the pETDuet vector was used to express the monooxygenase. With these data in hand, the co-expression of StyAB and LsADH from pETDuet and pCDFDuet-1, respectively, was tested in a 2 L bioreactor.

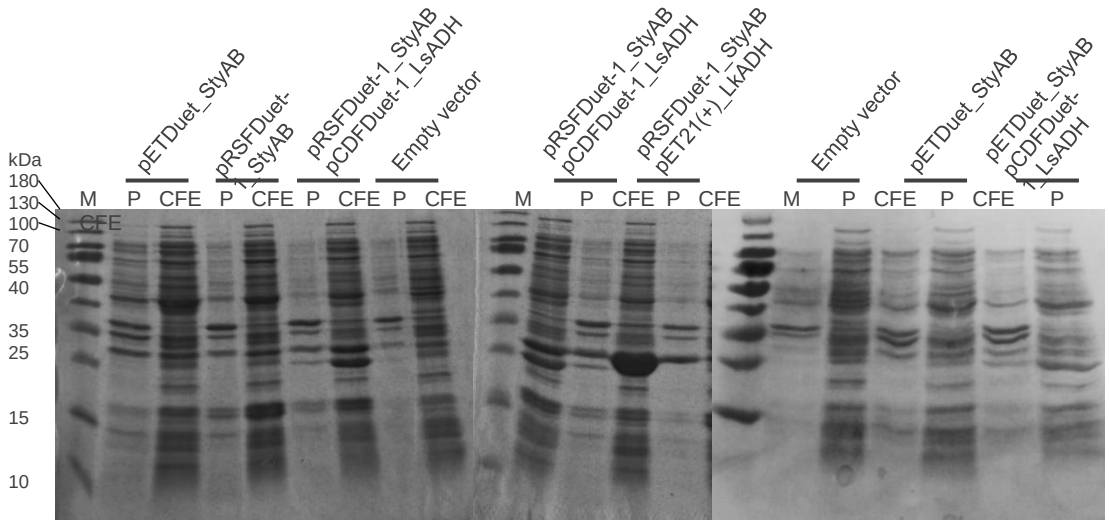


Figure 10. Exemplary gels showing the expression levels of StyA (45kDa) and StyB (not visible) and relative ADH (LkADH 26 kDa, LsADH 26.5 kDa) in different expression vectors. LkADH expresses significantly better but at the same time, it suppresses StyA expression.

The use of a bioreactor compared to a shake flask allows fine-tuning of several parameters. For example, it is possible to maintain a desired dissolved oxygen concentration (e.g. 30%) and setting the agitation speed and air input dependent on this value. This generally results in a gradual increase of the agitation speed directly dependent on the cellular concentration giving that biomass consumes more oxygen as

it grows. The gradual increase in agitation speed over time can be used as proof of the cells' growth and; it is also possible to maintain the pH constant to the desired level by titrating either a base or acid solution when the pH is higher or lower than the set value respectively. StyAB and LsADH were coexpressed in *E. coli* BL21 (DE3) at 20 °C using 0.2 mM of IPTG as determined in shake flask experiments (Table 18). In a typical experiment with *E. coli* BL21 (DE3), the dissolved oxygen is set constant to 30 %, agitation speed starts at 200 rpm and it can increase until 800 rpm to maintain the dO_2 level. 800 rpm is the maximum speed at which the bacteria can still withstand the mechanical stress. After this value is reached, the airflow volume, typically depending on the culture volume (e.g. 1.5 L/h for 1.8 L culture), starts to increase. Figure 11 shows the time course of two fermentations and in both cases, the agitation speed increased to maintain a dO_2 level of 30%. The agitation speed first reached a plateau and then decreased in correspondence of the stationary phase due to a slower cellular metabolism caused by nutrients depletion. In both runs, the maximum obtained OD_{600} was around 20. The activity was measured during the course of the fermentation to determine an optimal time to harvest the cells. The activity was measured by monitoring the product formation by GC after 5 min (as mentioned previously); in this case, cell-free extract was used instead of whole cells. Highest epoxidation activity was reached in early-stationary phase (Table 19, Fermentation A, t=12 h induction time). The activity appears to decrease as the bacteria enter the stationary phase and therefore the harvest should be planned accordingly to maximise cellular concentration and activity.

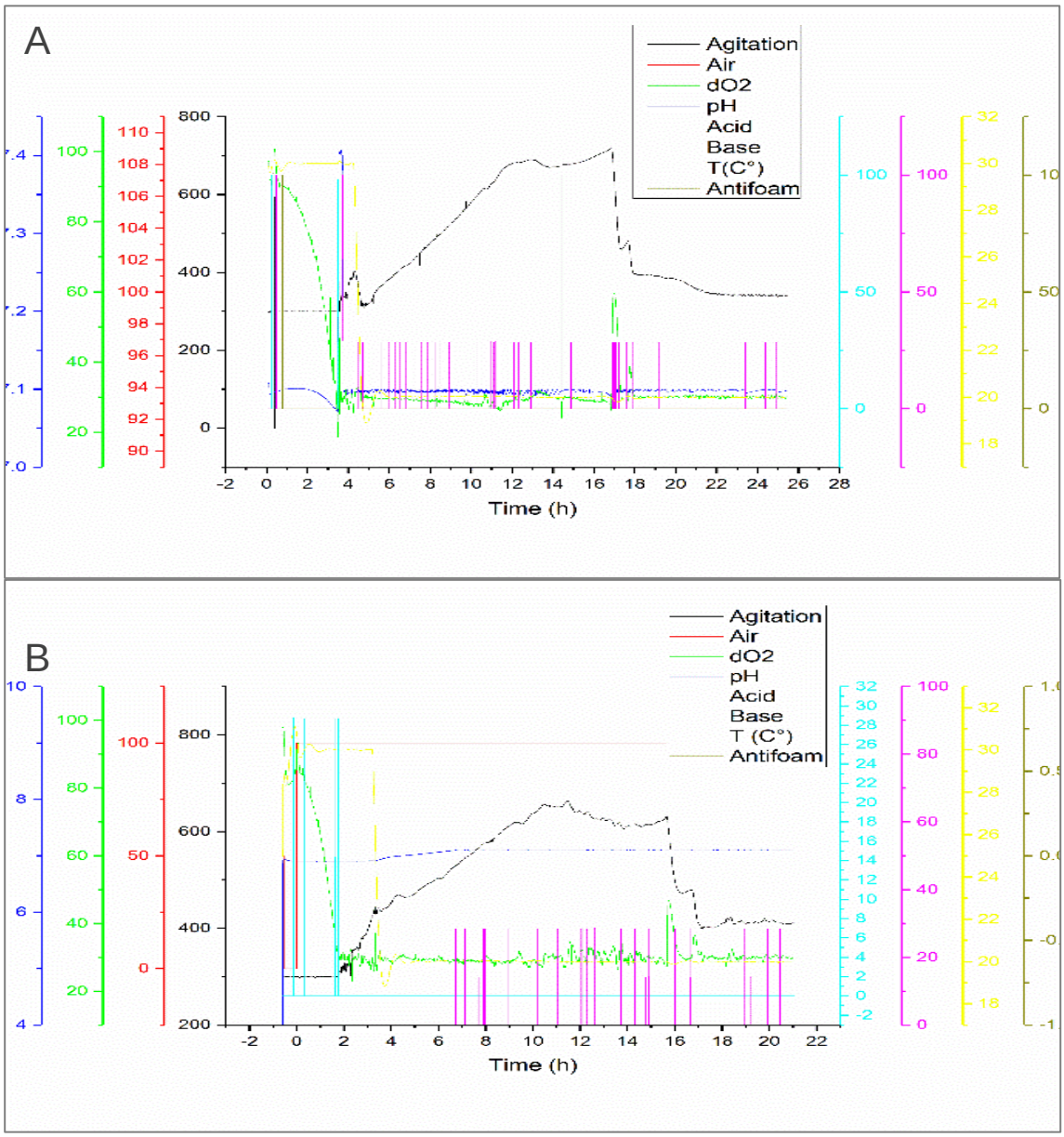


Figure 11. Fermentation graphs of the two expressions run. Important parameters are continuously recorded during the run. The green line (dissolved oxygen) stayed relatively stable around 30% while the stirring speed (black line) increased until the stationary phase (10-12h). Samples were taken at fixed times (Table 19) to measure the activity in the epoxidation of *trans*-2-heptene.

Table 19. Activity and final OD_{600} data corresponding the fermentation graphs shown in Figure 11. Samples were taken at different times after induction and the activity was measured by product formation over 5 min using CFE.

Fermentation	Induction time	OD_{600}	mU/mg (total protein content)
A	12h	18.72	5.97 ± 0.012
A	16h	22.8	4.83 ± 0.01
A	21h	20.2	4.48 ± 0.009
B	3h	4.5	4.29 ± 0.009
B	18h	19.0	5.59 ± 0.011

= (1U= 1 µmol/min)

3.1.1.3 Expression and purification of the cytochrome P450 monooxygenases CYP154H1 from *Thermobifida fusca* and CYP154C5 from *Nocardia farcinica* as well as respective electron transfer components

The styrene monooxygenase from *Rhodococcus sp.* ST-10 was not the only monooxygenase tested in the conversion of the selected substrates. Two cytochrome P450s, CYP154H1 from *Thermobifida fusca* and CYP154C5 from *Nocardia farcinica*, were expressed and purified together with putidaredoxin (Pdx) and putidaredoxin reductase (PdR) from *Pseudomonas putida* as electron transfer components using an established protocol.[159,181] *E. coli* C43 (DE3) was used as expression host as demonstrated previously and additionally, each cytochrome P450 was coexpressed with PdR and Pdx to test in whole cell biocatalysis experiment. Both monooxygenases carried an N-terminal His-tag, hence the purification was carried out via immobilized metal affinity chromatography (IMAC) using a Ni-NTA column connected to an FPLC system (Äkta). The chromatogram and SDS-PAGE gel of the purification of CYP154C5 are shown in Figure 12 as an example purification. The green line shows the progress of the gradient used for the elution. For the gradient, a solution with a high imidazole concentration (500 mM) is mixed in increasing quantities with the equilibration buffer (30 mM imidazole). Imidazole competes with the histidines in the his-tag of the target protein in binding the metal in the column causing the protein to elute. Non-target proteins elute directly after the column is loaded, as visible in the large peak at the beginning of the chromatogram in Figure 12 (blue line). The target protein started to elute from the column at approximatively 170 mM imidazole concentration (30% of the gradient).

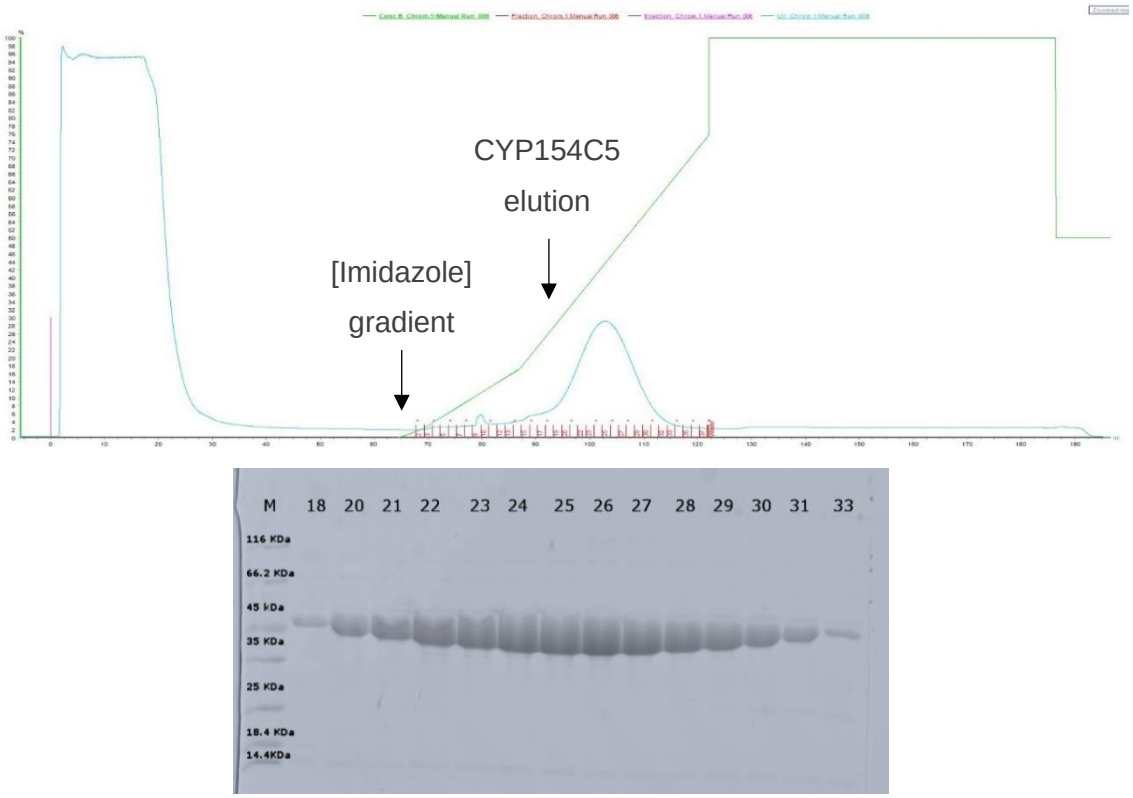


Figure 12. Typical purification chromatogram of a His-tagged protein where almost the total content of non-tagged protein in the cell-free extract elutes in the first fractions while the target protein starts eluting only after the imidazole concertation increases (UV signal, light blue line). The chromatogram shows the imidazole concentration in % (green line) of the highest concentrated buffer relative to the solution used to equilibrate the column. In the gel, only the fractions containing the target protein are shown. All fractions appear to be clean except a small band visible around 14.4 kDa.

Only the fractions of the elution of the target protein (~45 kDa) are shown in the gel in Figure 12). All fractions appeared clean except for a small band visible around 14.4 kDa. It was possible to follow the progression of the purification of all cytochromes P450 due to the characteristic red-brown colour of the P450s caused by the heme prosthetic group. Final pure protein yields were for CYP154H1 and CYP154C5 52 and 43 mg/L (mg per litre of culture), respectively. The concentration of both cytochromes was determined by CO-difference spectroscopy measuring the absorbance maximum of CO-bound CYP154C5 or CYP154H1 at 450 nm (ε_{450}= 91 mM^{-1} cm^{-1}).[169] In contrast, the electron transfer components PdX and PdR were expressed without any affinity tag and they required a different purification strategy. PdR purification was carried out in two steps (I) anion exchange (AIXC) and (II) hydrophobic interaction (HIC) chromatography, while for Pdx purification, the strategy was to combine anion exchange chromatography with an ultrafiltration step using a 30 kDa cut-off membrane to remove larger proteins. Figure 13 shows the chromatograms of PdR purification, in the first step a Q-Sepharose FF ion exchange column was used and the progression of the purification was followed by the characteristic yellow colour that indicates the presence of the FAD cofactor. A good

amount of the *E. coli* proteins was eliminated in the first step as visible in the first chromatogram in Figure 13. In the second step, a Phenyl-Sepharose FF hydrophobic interaction matrix was used to complete the purification and obtain a clean single band in the SDS-PAGE gel (Figure 13).

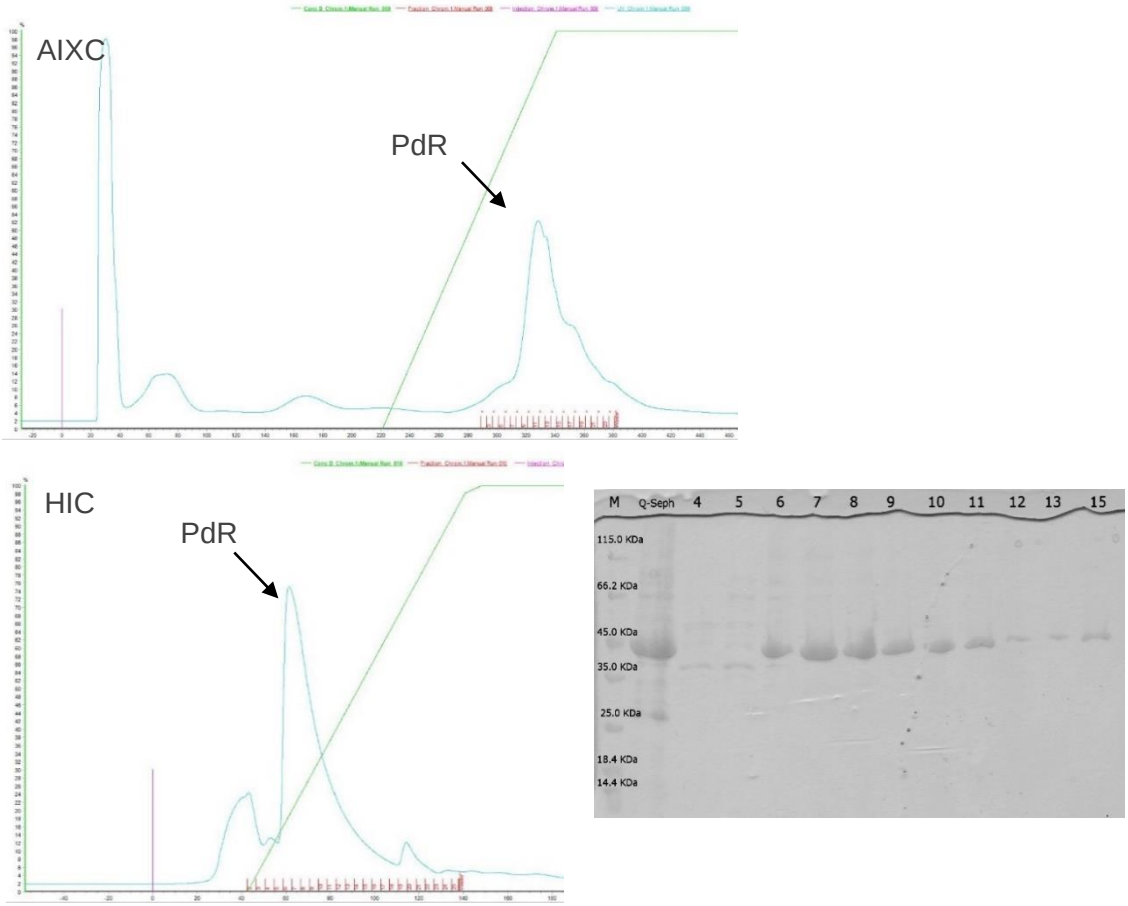


Figure 13. Two steps purification of PdR. The first chromatogram shows the anion exchange purification (AIXC, first step) where PdR started to elute at 350 mM KCl (70% of the gradient). The second chromatogram (HIC) shows the hydrophobic interaction purification (second step) where PdR started to elute at about 1.12 M KCl (15% of the gradient). The gel picture shows that the second step is needed to afford the reductase in pure form. After the AIXC step, some impurities from the cellular proteome are still present as shown in the line “Q-seph”.

The concentration of pure PdR was determined by measuring the absorbance at 454 nm (ε_{455}=10.0 mM^{-1} cm^{-1}).[170] After concentration by ultrafiltration with a 30 kDa cut-off membrane, the purification yielded 82 mg/L culture. As mentioned before, Pdx was purified in two steps whereas the first purification was the same as for PdR, anion exchange chromatography using Q-Sepharose FF as matrix. The second step was ultrafiltration using a 30 kDa cut-off membrane. This step was used to filter out the larger proteins, while Pdx having a molecular weight around 11 kDa could pass through. The flow-through can then be concentrated with a 10 kDa cut-off membrane. Figure 14 shows that many of the impurities were eliminated in the first purification step and several peaks

preceding the target protein can be seen in the chromatogram. Additionally, the impurities in the latter appeared to be significantly less when comparing the bands “CFE” with the fractions containing Pdx in the gel. Nevertheless, it is only after the ultracentrifugation that it was possible to yield the redoxin in pure form yielding 7 mg/ L culture. As for the previous cases, the purification progress could be followed by the typical brown colour of proteins containing an iron-sulphur cluster. The concentration of pure Pdx was determined at 455 nm (ε_{455}=10.4 mM^{-1} cm^{-1}).[170]

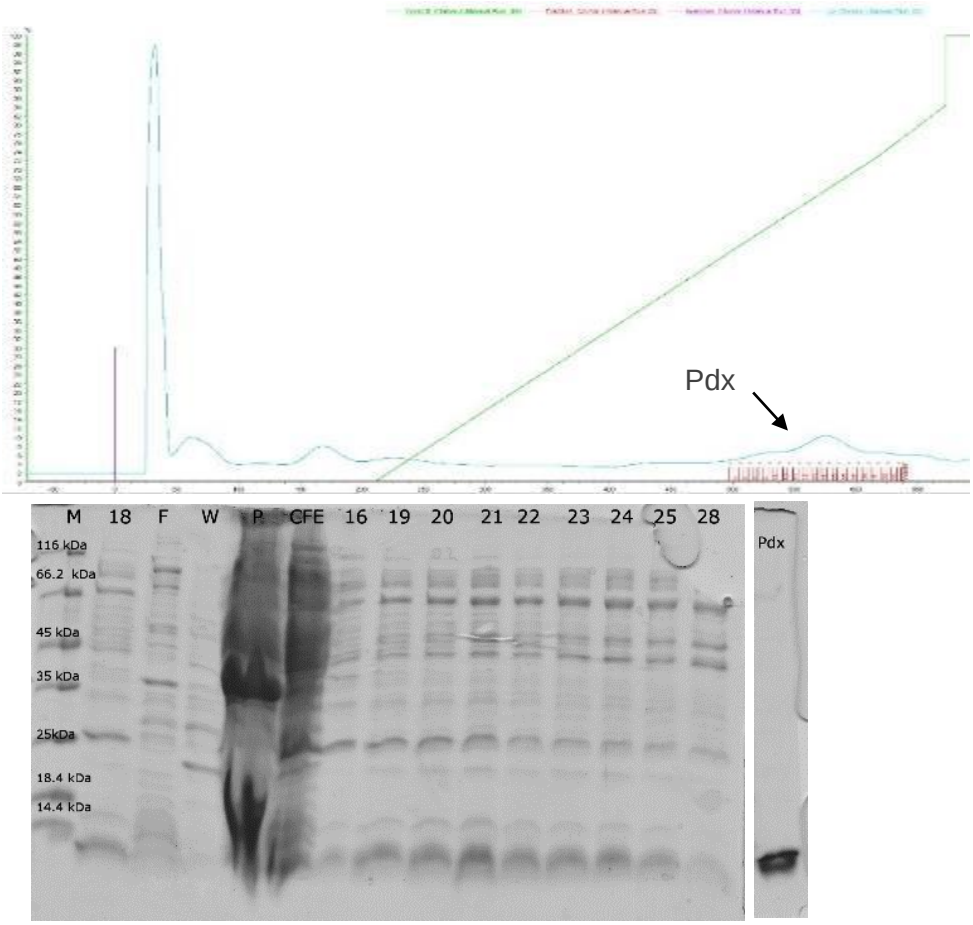


Figure 14. Two steps purification of Pdx. The chromatogram shows the first purification step using the anion exchange column (AIXC), Pdx started the elution around 300 mM KCl (60% of the gradient). The gel shows the fractions after the first step purification where several bands at higher kDa than Pdx can be seen. The last line “Pdx” shows the pure protein after filtration through a 30 kDa cut-off membrane and after concentration using a 10 kDa cut-off membrane.

3.1.1.4 Screening of enzymatic and chemical epoxidation catalysts

All acyclic alkenes used in this work are prochiral molecules that, upon epoxidation, result in two different enantiomers of the corresponding epoxides. Hence, our strategy for the production of enantiopure amino alcohols from these acyclic alkenes was to control the stereocenters already in the first step (epoxidation) by synthesizing enantiopure epoxides. For this, a set of different enzymatic and chemical catalysts were tested for their ability to yield enantiopure epoxides using *trans*-2-heptene as a model substrate. The results of all active catalysts are shown in Figure 15. Unfortunately, none of the cytochromes P450 described in the previous section displayed activity towards the selected substrates. Hence, no further tests were done using these enzymes. Notably, only the styrene monooxygenase and the Shi epoxidation catalyst displayed significant activity. Conveniently, the styrene monooxygenase synthesized the (*S,S*)-epoxide almost exclusively with an enantiomeric excess higher than 95%, whereas the Shi diketal epoxidation catalyst produced the (*R,R*)-enantiomer with an enantiomeric excess around 90%. Hence, two catalysts could be identified yielding opposite epoxide enantiomers in the first step. Ultimately, this will also enable the synthesis of opposite enantiomers of the final aminoalcohol products.

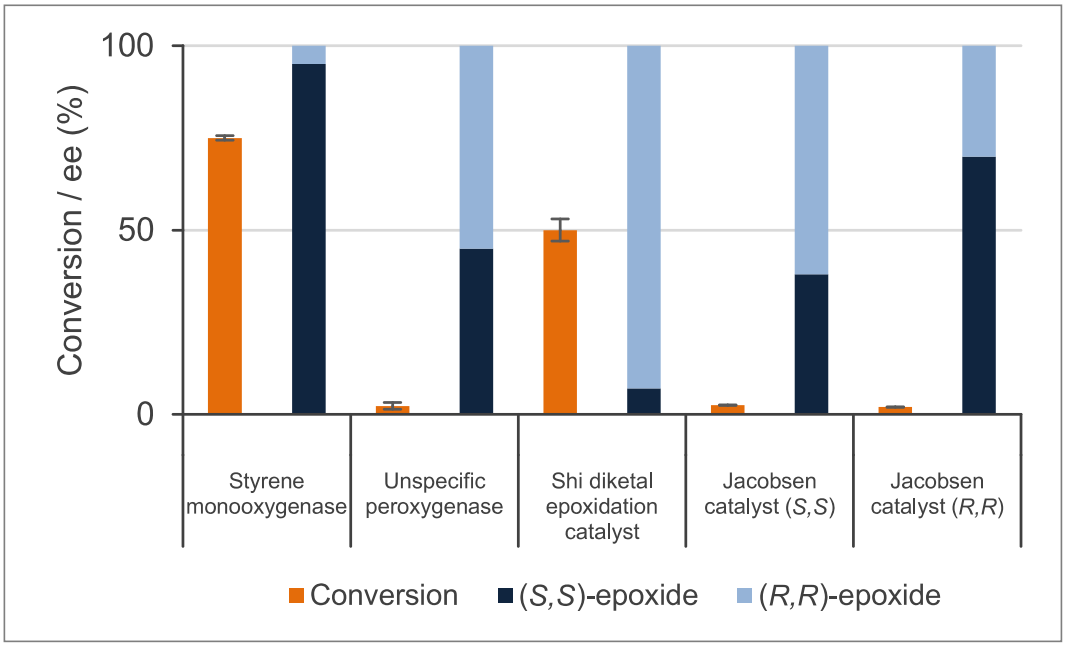


Figure 15. First screening for the epoxidation of *trans*-2-heptene catalysed by different enzymatic or chemical catalysts. First two entries are enzymatic catalysis and last three entries are chemical catalysts.

3.1.1.5 Optimization of styrene monooxygenase-catalysed epoxidations on small scale

The styrene monooxygenase is a two-component system that requires NADH and $FADH_2$ to provide the electrons necessary to activate molecular oxygen (Scheme 9). Hence, a cofactor regeneration system is desirable to avoid the addition of cofactors in stoichiometric amounts. For the experiments in Figure 16, two possible strategies for cofactor regeneration were used. If whole cells (WC) are used, glucose can be added to fuel cellular metabolism that can regenerate NADH. However, when cell-free extract (CFE) is used the cellular metabolism is not working anymore. Thus, an enzyme oxidizing a co-substrate is used to reduce NAD^+ to NADH. For these reactions, the formate dehydrogenase from *Candida boidinii* was used to oxidise formate for cofactor regeneration.

The setup of an efficient epoxidation using a monooxygenase is a complex task requiring the analysis of many variables. At first, the use of whole cells or cell-free extract in combination with a balloon of pure oxygen or just leaving sufficient air in the headspace was investigated (No O2 entries, Figure 16). Both model substrates *trans*-2-hexene and *trans*-2-heptene (**1** and **2**) were analysed using similar conditions and in both cases, the highest conversion in reactions using 5 mM substrate was obtained when the styrene monooxygenase was used as cell-free extract coupled with the formate dehydrogenase from *Candida boidinii* added as cofactor regeneration system (Figure 16). Furthermore, the use of pure oxygen showed an additional beneficial effect and almost 100% conversion of *trans*-2-heptene could be achieved within 1 h.

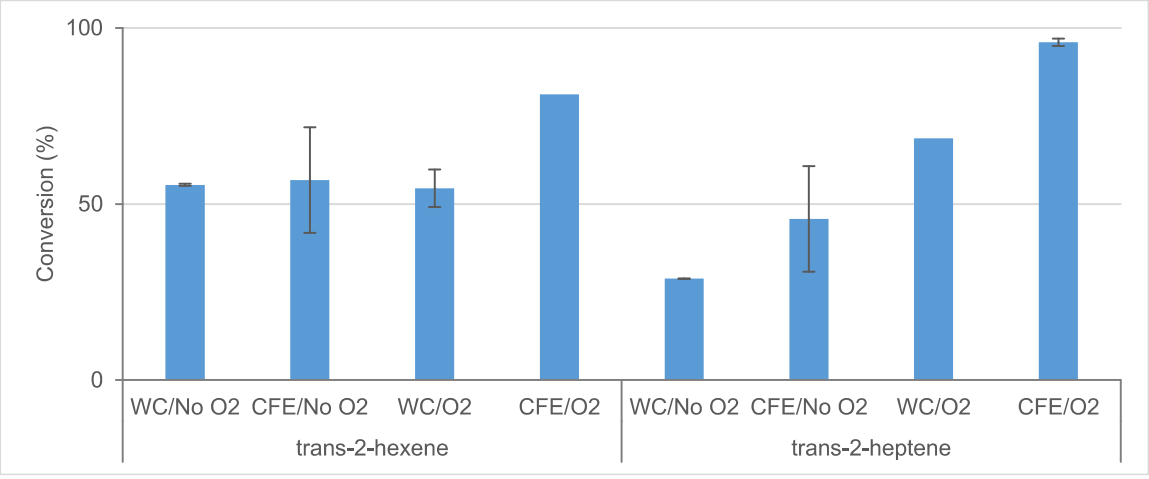


Figure 16. Comparison of the performance of StyAB when used as whole-cell biocatalyst (WC, 60 g/L wet cell weight) or as cell-free extract (CFE, 9.6 g/L total protein content). FDH (0.5 U/mL) was used as cofactor regenerating enzyme coupled with CFE, while glucose was added for cofactor regeneration in the WC system. Oxygen was either supplied via the air in the headspace ("No O_2") or using pure oxygen ("O_2") provided in a balloon.

Another recently developed strategy is to simplify the reaction system by removing StyB and cofactor regeneration using a molecule that can transfer electrons directly to the FAD thanks to its specific redox potential (as mentioned in section 1.4, Scheme 10). One of these molecules, 1-benzyl-1,4-dihydronicotinamide (BNAH), was proven to be a good non-enzymatic regeneration system for the styrene monooxygenase from *Rhodococcus opacus* 1CP.[91] Hence, different concentrations of BNAH and FAD were tested to investigate the feasibility of this system for the epoxidation of *trans*-2-heptane as model substrate. In Figure 17, it is shown that the use of 2 equivalents of BNAH (with respect to the substrate concentration) efficiently provided electrons for StyA in order to perform the epoxidation of 5 mM substrate. In contrast, varying the FAD concentration had no significant effect. In both cases, it was possible to reach over 90% conversion in 1h. However, when the BNAH concentration was doubled while keeping the same substrate concentration, the conversion dropped to 13% indicating a negative influence on the epoxidation by BNAH. A similar effect is obtained when the substrate concentration is increased to 10 mM, restoring the 1:2 ratio of substrate:BNAH, as the conversion reached only 21%. This made BNAH not an ideal candidate for a possible scale-up, as significantly higher substrate loadings are required.

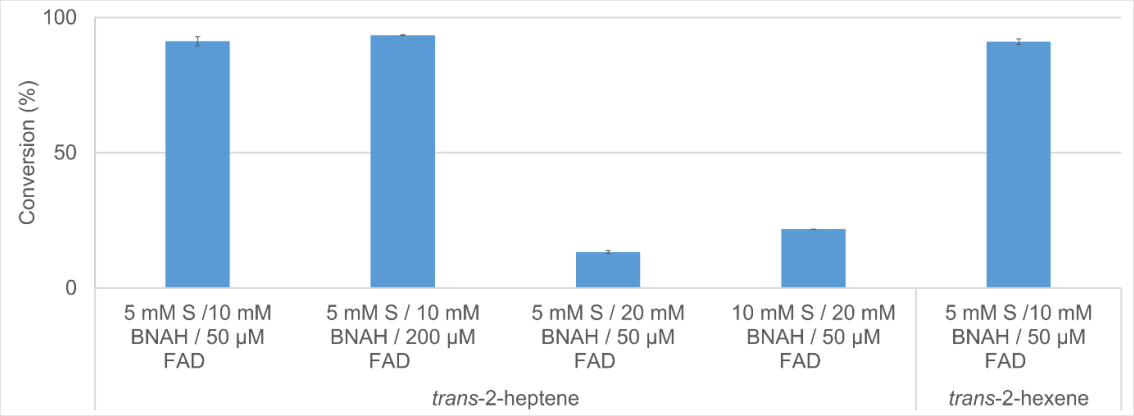


Figure 17. Conversions obtained after 1h in reactions using different substrate, BNAH and FAD concentrations and cell-free extract harbouring StyA (9.5 g/L, total protein concentration). *Trans*-2-heptene was first used to find optimal conditions (5 mM substrate, 10 mM BNAH, 50 µM FAD) and then they were used in the conversion of *trans*-2-hexene where similar conversions could be achieved.

The availability of O_2 is a critical factor that needs to be optimised when working with monooxygenases. If glucose is used for the cofactor regeneration, the cellular metabolism consumes additional oxygen lowering its overall availability for enzyme activity. Therefore, a good compromise must be found to maximize the availability of O_2 and reduced cofactor for the enzyme. Figure 18 shows the activity values measured with

different concentrations of whole cells harbouring StyAB using 2.5 mM. All activities were determined by measuring the product formation over 5 min.[78] It is visible that the highest activity was reached when only 15 g/L (wet cells weight) of biomass was present. Doubling the concentration of biocatalyst caused a 10% decrease in the activity and further increased cellular concentration did not show a significant effect. It appears that there is an optimal cellular concentration to ensure the best use of O_2 between cellular metabolism and enzyme activity. In fact, when the biomass concentration is too high it consumes most of the available oxygen in solution leading to lower activity.

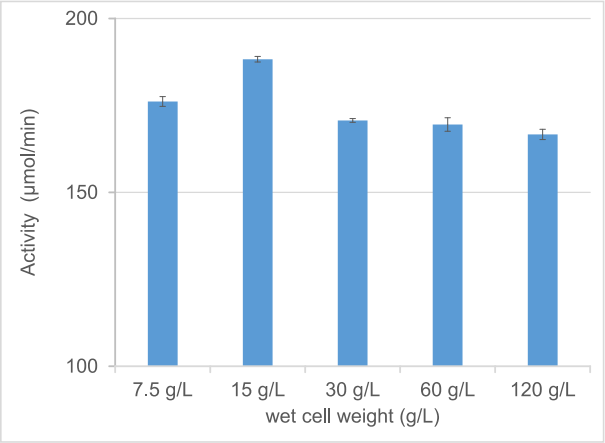


Figure 18. Activity measured for StyAB used as whole cell biocatalyst using glucose for cofactor regeneration. The highest activity in the conversion of 2.5 mM *trans*-2-heptene is reached with 15 g/L wet cell weight.

After the optimization of cellular concentration, the stability of StyA was tested over time. BNAH was used to reduce FAD to be able to exclude a possible deactivation of StyB. The assay was carried out with CFE containing StyA at 30 °C with magnetic stirring in potassium phosphate buffer. At different time points, samples were taken and the activity was measured and compared with the fresh sample's activity. As can be seen in Figure 19, almost complete loss of activity was observed after 6 h. This might indicate low stability of StyA; however, it could be also due to mechanical stress caused by the magnetic stirrer.

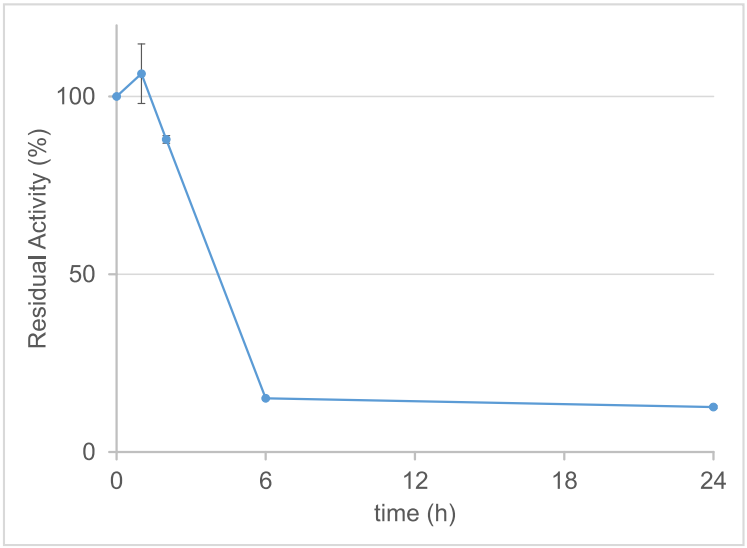
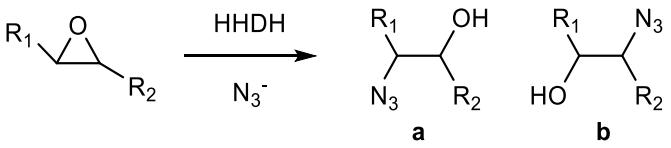


Figure 19. Residual activity of StyA after incubation at 30 °C for different time points. After 6h most of the activity is lost.

3.1.2 Ring-opening of di-substituted epoxides

After successful synthesis of di-substituted epoxides, a set of 22 HHDHs was tested for their activity in the ring opening of these epoxides using azide as nucleophile. Until recently, only terminal epoxides had been reported to be accepted as substrates by HHDHs.[111] Only in one paper by Hasnaoui-Dijoux et al., it is briefly noted that HheC was observed to catalyse the ring-opening of 2,3-epoxyheptane with nitrite as nucleophile but without experimental detail.[182]

Within our study, several di-substituted epoxides were chosen to further explore the substrate scope of our newly identified HHDHs. Scheme 17 shows that two different regioisomers can be formed in the azidolysis of a di-substituted epoxide depending on the site of nucleophilic attack.



Scheme 17. General scheme for the HHDH-catalysed azidolysis of di-substituted epoxides resulting in the formation of two possible regioisomeric azidoalcohols.

As described in the aim of the project (section 1.8), a regioselective epoxide ring opening was required as the second step in order to set up a viable cascade for the production

of enantiopure amino alcohols from alkenes. Therefore, the major goal was to identify regioselective HHDHs with high activity towards the target substrates.

3.1.2.1 Expression and purification of halohydrin dehalogenases in shake flasks

Expression and purification of all HHDHs were performed according to optimized protocols previously reported by Koopmeiners *et al.*[167] All HHDHs were produced in *E. coli* BL21 (DE3) or C43 (DE3) transformed with pET28a(+) vector containing the respective *hhdh* gene. All purifications were performed via immobilized metal affinity chromatography (IMAC) using a Ni-NTA column connected to an FPLC system (Äkta) making use of the enzyme´s N-terminal hexa-histidine tag. An example of a typical chromatogram and SDS-PAGE gel of the purification of HheE2 is shown in Figure 20. Elution fractions containing the target protein are shown in the gel (Figure 20). All fractions appeared to be clean except for a small band visible around 14.4 kDa. Purified HHDHs were usually obtained in yields of 2-26 mg/L culture.[105,167]

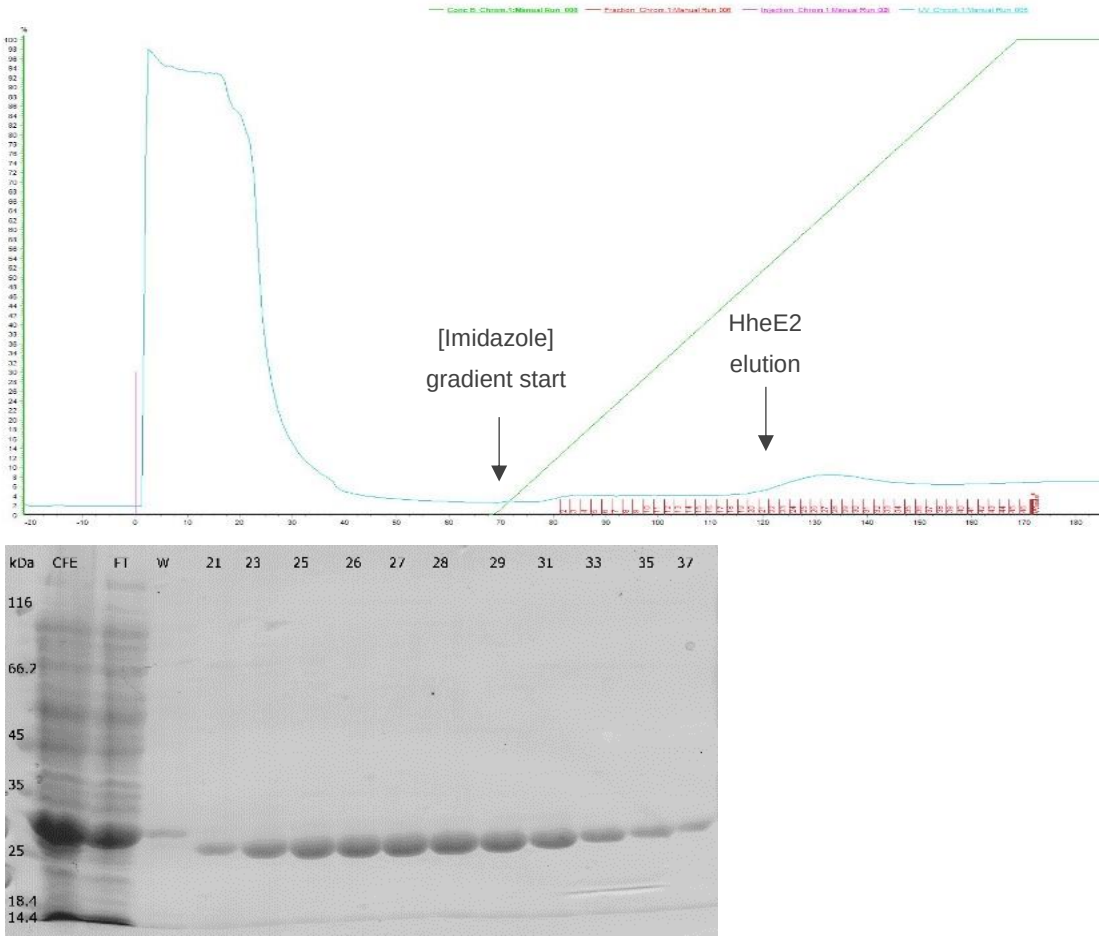


Figure 20. Classic purification chromatogram of His-tagged HHDH, HheE2 in this case. All HHDHs started to elute around 300 mM of imidazole concentration. The gel only shows the fractions containing the target protein. All fractions appear to be fairly pure.

Expression of HheE and HheE5 was also tested in a 2L bioreactor. Both enzymes were produced in *E. coli* BL21 (DE3) at 20 °C using 0.1 mM of IPTG as inducer, and using the same protocol adopted for the styrene monooxygenase (Section 3.1.1.2). As seen in Figure 11, the HHDHs' fermentations showed a similar graph (data not shown), after overnight expression samples were taken to compare the activity with shake flasks fermentations (Table 20).

Table 20. Specific activities of HheE and HheE5 after expression in the 2 L bioreactor. Samples were taken after overnight expression and the activity of the CFE was measured by using the spectrophotometric pH assay (Methods, Enzyme assays). Specific activity values are based on total protein content.

Enzyme	Cultivation	Final OD_{600}	Specific activity (mU/mg)
HheE	Fermentation 1	10.3	3.98
HheE	Fermentation 2	13.4	6.6
HheE	Shake flask	6.1	7
HheE5	Fermentation 1	9.74	26.2
HheE5	Fermentation 2	15	11
HheE5	Shake flask	9	44.5

= (1U= 1 µmol/min)

The values in Table 20 show that HHDHs produced in the bioreactor have lower specific activity than HHDHs produced in shake flasks. HheE5, especially, reached up to 4 fold less activity compared to shake flasks experiments. Possibly, HheE5 requires shorter expression times as the main difference between the samples is the significantly higher final OD_{600}. This is compatible with Koopmeiners *et al.*, as HheE5 was typically harvested after 7 h from induction at 30 °C.[167] In a bioreactor, longer times are needed to reach sufficient OD_{600} to start induction and production would end a later time than the one allowed by the safety measures. HheE's activity, in contrast, does not seem to be dependent on the induction time.

3.1.2.2 Selective ring opening of di-substituted epoxides

A subset of 22 HHDHs from our panel, containing enzymes from all six phylogenetic HHDH subtypes, was tested in bioconversions of different di-substituted epoxides (**21-25**) using azide as exemplary nucleophile (Scheme 17). The majority of the tested HHDHs (17 out of 22) displayed activity in the conversion of the two aliphatic substrates *trans*-2,3-epoxyhexane (**21**) and *trans*-2,3-epoxyheptane (**22**), affording the corresponding azidoalcohols in varying yields. In contrast, significantly fewer HHDHs converted the sterically more demanding epoxides *trans*-4,5-epoxyhexan-1-ol (**23**),

trans-3,4-epoxyheptane (**24**) and *trans*-β-methylstyrene oxide (**25**). HheG from *Ilumatobacter coccineus* was active on all five epoxide substrates yielding highest conversion for substrates **21**, **23**, **24** and **25** among all tested HHDHs. Significant conversion of **25** by HheG was unexpected as the enzyme was previously shown to exhibit only very little activity on styrene oxide, a structurally very similar but terminal epoxide.[167] HheA2 and HheB2 hardly showed any product formation with any of the substrates, which is in agreement with previous findings.[111,182] Interestingly, the addition of a terminal hydroxyl group on the opposite end of the substrate compared to the epoxide ring completely suppressed activity of all tested HHDHs that displayed activity on the corresponding non-hydroxylated epoxide **21.** As shown in Figure 21 (gold-yellow bar), HheG was the only exception still able to convert *trans*-4,5-epoxyhexan-1-ol (**23**) with a significant yield, notably HheG2 showed no conversion at all, albeit both enzymes share 74% sequence identity at the protein level.[105] Comparing substrates **22** and **24**, which differ in the position of their epoxide ring, only a reduced number of HHDHs was still active on *trans*-3,4-epoxyheptane (**24**). Enzymes HheA3, HheD, HheD2, HheD3 and HheD5, as well as HheG2 gave low to moderate conversions, whereas HheG converted **24** completely within 15 h reaction time (Figure 21, Green).

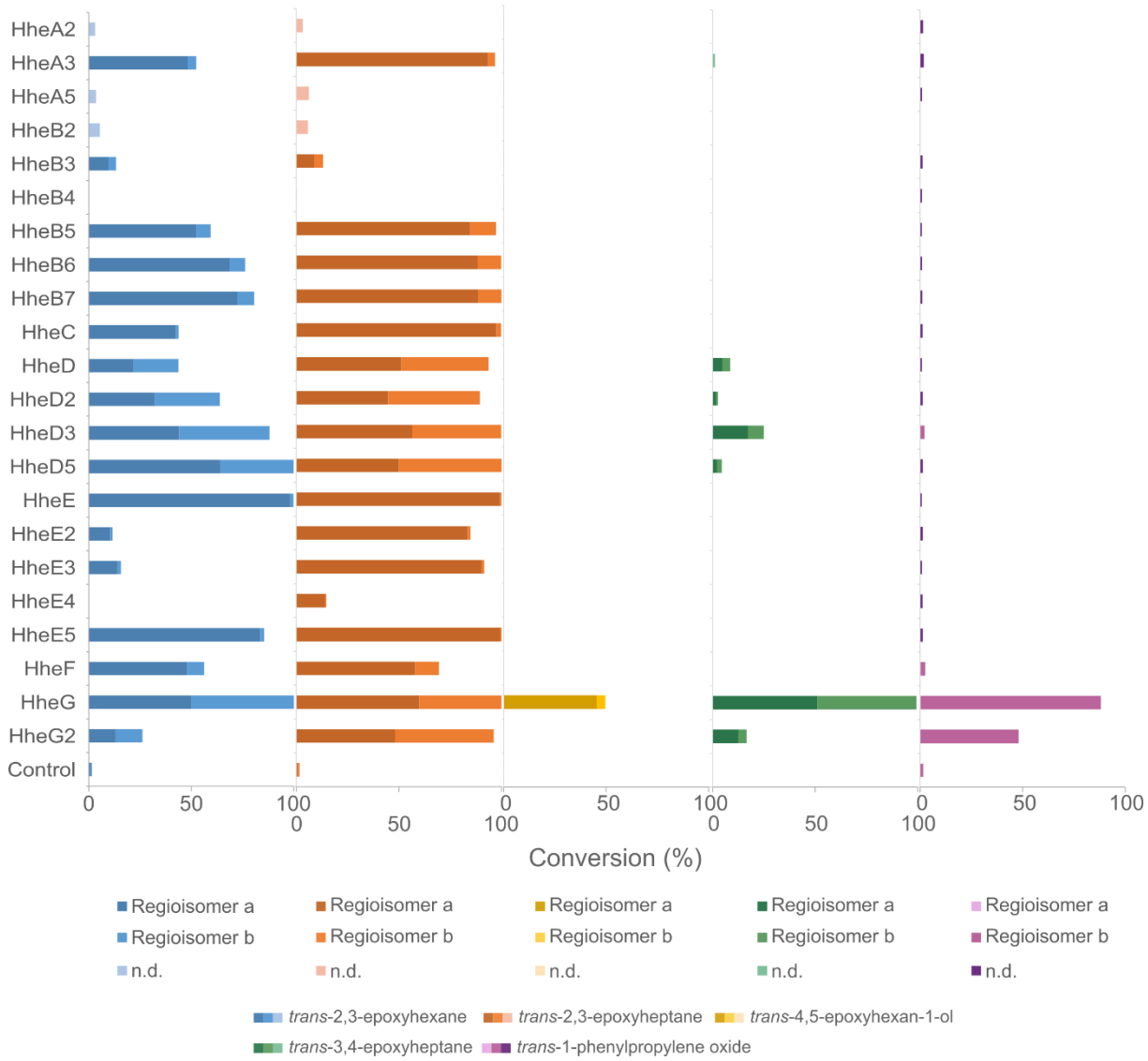
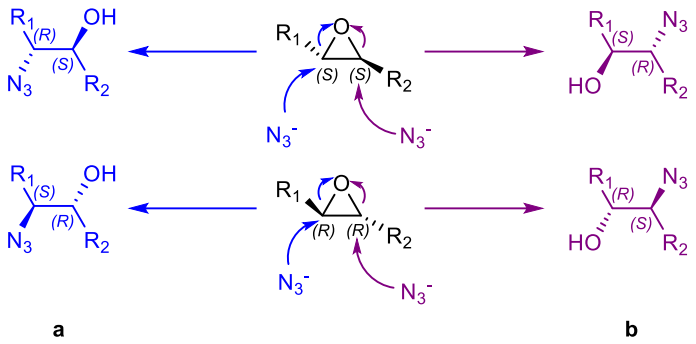


Figure 21. Conversions and ratios of regioisomeric products **a** and **b** obtained in biocatalytic reactions of substrates **21-25** with HHDHs using azide as the nucleophile. The ratios of regioisomeric products **a** and **b** are represented in different colours within the bar representing the conversion.

The vicinally di-substituted epoxide substrates **21-25** each contain two stereocenters. Depending on the site of the nucleophilic attack and the stereoconfiguration of the substrate during HHDH-catalysed epoxide ring-opening, two regioisomers and their respective pairs of enantiomers can be attained (Scheme 18). In agreement with our cascade strategy for selective synthesis of enantiopure aminoalcohols from non-terminal alkenes, the regioselectivity of the catalyst in the second step is of paramount importance. Thus, the regioselectivity of active HHDHs in the conversion of substrates **21-25** was studied. Most enzymes formed both regioisomeric products **a** and **b** in varying ratios (Figure 21). HheA3, HheC and all tested enzymes from the E subfamily exhibited high regioselectivity for substrates **21** and **22**, giving 2-azido-3-hydroxyhexane (**31a**) and 2-azido-3-hydroxyheptane (**32a**) almost exclusively. In contrast, enzymes from HHDH

subfamily D as well as HheG and HheG2 formed both regioisomers (**31a** and **31b**, as well as **32a** and **32b**) in roughly equal amounts. The same was observed for chemical background azidolysis of **21** and **22** ("Control" in Figure 21). Compared to the conversion of **21**, the terminal hydroxyl group present in **23** had a positive effect on HheG's regioselectivity. The enzyme produced 5-azidohexane-1,4-diol (**33a**) with high preference (Figure 21).



Scheme 18. Possible regio- and stereoisomers in the azidolysis of vicinally di-substituted trans-epoxides **21–25**. The site of attack (α: less sterically hindered, shown in blue; β: sterically more hindered, shown in plum) defines the regioisomers obtained, while the enantioselectivity of the enzyme and the S_N2-type mechanism define the absolute configuration found in each set of regioisomers. Hence, a total of two regioisomers **a** and **b**, each with their respective (*R,S*)- and (*S,R*)-enantiomers, can be obtained. Adapted from Calderini *et al.*[183]

None of the active enzymes seemed to display high regioselectivity in the conversion of **24** as in all cases both azido alcohol regioisomers **34a** and **34b** were formed. This might not be surprising since the substituents of the epoxide ring are more similar in size, compared to (**22**). In accordance with literature,[184,185] none of the tested enzymes exhibited preference towards a nucleophilic attack on the sterically more hindered carbon atom, except for substrate **25**. In this specific case, the nucleophilic attack on the benzylic carbon is favoured due to the electronic resonance effect of the aromatic substituent. Here, HheG and HheG2 gave azido alcohol **35b** almost exclusively, whereas HheD3 and HheF formed also small amounts of regioisomer **35a**. Chemical azidolysis of epoxide **25** afforded **35b** almost exclusively as well. In this specific case, nucleophilic attack on the benzylic carbon is favoured due to the electronic resonance effect of the aromatic substituent.

Furthermore, the enantioselectivity of all active enzymes in the ring-opening of racemic epoxides **21** and **22** was analysed. For this, product enantiomeric excesses (ee_P) and corresponding apparent E-values for each of the two possible product regioisomers were determined. As a result, most azidoalcohols were formed with low to moderate enantioselectivity (Table 21). Several HHDHs, however, displayed good

enantioselectivity (E> 15) in the formation of azido alcohol regioisomers **21b** and **22b**. Hence, nucleophilic attack on the unfavoured carbon atom occurred with higher enantioselectivity. Interestingly, most azido alcohol products in the conversion of **21** and **22** were preferentially formed in 2*S*,3*R*-configuration. This indicates that epoxides **21** and **22** with *R,R*-configuration were preferentially attacked at the sterically less hindered carbon atom (C2), whereas regioisomers **21b** and **22b** were preferentially formed by nucleophilic attack of (*S,S*)-**21** and (*S,S*)-**22** at the sterically more hindered carbon atom (C3). In contrast, nucleophilic attack of azide at C2 in (*S,S*)-**21** and (*S,S*)-**22** yields azidoalcohols **21a** and **22a** in 2*R*,3*S*-configuration, as found for HheA3, HheE2, HheF and HheG2 with epoxide **21** and HheE4 and HheF with epoxide **22** (Table 21). A detailed analysis of the enzymes' regiopreferences in the azidolysis of *R,R*- and *S,S*-enantiomers of epoxides **21** and **22** is given in Supporting table 1 the appendix. Interestingly, D-type HHDHs displayed a slightly higher preference for nucleophilic attack of (*S,S*)-**21** and (*S,S*)-**22** at the sterically more hindered carbon atom, whereas most other HHDHs exhibited a clear preference for nucleophilic attack of the sterically less hindered carbon atom independent of the substrate's stereoconfiguration. Additionally, the enantioselectivity of active HHDHs in the ring-opening of *rac*-**25** was determined and results are given in Supporting table 2 in the appendix. No enantioselectivity was observed in the chemical epoxide ring opening of **21**, **22** and **25** on preparative scale with azide as nucleophile (data not shown).

Table 21. Conversions (C), enantiomeric excesses (ee_P) and calculated apparent enantiomeric ratios (E^{app}) of regioisomeric azidoalcohols **21a**, **b** and **22a**, **b** formed in the HHDH-catalysed azidolysis of epoxides **21** and **22**. The absolute configuration of the preferentially formed enantiomer of each regioisomer is given in parentheses. Table taken from Calderini *et al.*[183]

	6a			**6b**			**7a**			**7b**		
HHDH	C [%]	ee_P [%]	E^{app}	C [%]	ee_P [%]	E^{app}	C [%]	ee_P [%]	E^{app}	C [%]	ee_P [%]	E^{app}
HheA3	48	47	4.2 (*R*,*S*)	4.1	46	2.8 (*S*,*R*)	93	0.7	1.0 (*S*,*R*)	3.4	25.5	1.7 (*S*,*R*)
HheB3	9.7	35	2.2 (*S*,*R*)	3.6	66	5.0 (*S*,*R*)	8.7	30	1.9 (*S*,*R*)	4.3	77	7.7 (*S*,*R*)
HheB5	52	25	2.1 (*S*,*R*)	7.1	86	14 (*S*,*R*)	85	24	1.6 (*S*,*R*)	13	95	39 (*S*,*R*)
HheB6	69	24	2.6 (*S*,*R*)	7.6	89	18 (*S*,*R*)	88	16	1.4 (*S*,*R*)	11	93	28 (*S*,*R*)
HheB7	73	24	2.9 (*S*,*R*)	8.1	93	30 (*S*,*R*)	89	17	1.4 (*S*,*R*)	11	94	32 (*S*,*R*)
HheC	43	34	2.6 (*S*,*R*)	1.3	n.d.[b]	n.d.[b]	97	4.6	1.1 (*S*,*R*)	2.3	90	19 (*S*,*R*)
HheD	22	28	1.9 (*S*,*R*)	22	86	17 (*S*,*R*)	51	60	4.0 (*S*,*R*)	43	87	14 (*S*,*R*)
HheD2	32	22	1.7 (*S*,*R*)	32	84	17 (*S*,*R*)	45	57	3.7 (*S*,*R*)	45	78	8.1 (*S*,*R*)

HheD3	44	65	7.7 (*S*,*R*)	44	43	3.4 (*S*,*R*)	56	76	7.4 (*S*,*R*)	43	80	9.0 (*S*,*R*)
HheD5	64	46	6.5 (*S*,*R*)	36	44	3.2 (*S*,*R*)	50	13	1.3 (*S*,*R*)	50	11	1.2 (*S*,*R*)
HheE	98	4.9	1.1 (*S*,*R*)	1.7	n.d.[b]	n.d.[b]	99	3.5	1.1 (*S*,*R*)	0.8	n.d.[b]	n.d.[b]
HheE2	10	71	6.4 (*R*,*S*)	1.1	n.d.[b]	n.d.[b]	83	11	1.2 (*S*,*R*)	1.5	n.d.[b]	n.d.[b]
HheE3	14	4.5	1.1 (*R*,*S*)	1.8	n.d.[b]	n.d.[b]	90	2.3	1.0 (*S*,*R*)	1.4	n.d.[b]	n.d.[b]
HheE4	<0.1	n.d.[b]	n.d.[b]	n.d.[b]	n.d.[b]	n.d.[b]	14	33	2.0 (*R*,*S*)	0.5	n.d.[b]	n.d.[b]
HheE5	84	13	2.2 (*S*,*R*)	1.8	n.d.[b]	n.d.[b]	99	3.5	1.1 (*S*,*R*)	0.6	n.d.[b]	n.d.[b]
HheF	48	54	5.4 (*R*,*S*)	8.0	32	2.0 (*S*,*R*)	58	32	1.9 (*R*,*S*)	12	19	1.5 (*S*,*R*)
HheG	50	16	1.6 (*S*,*R*)	50	15	1.5 (*S*,*R*)	60	63	4.4 (*S*,*R*)	40	89	17 (*S*,*R*)
HheG2	13	54	3.6 (*R*,*S*)	13	67	5.6 (*S*,*R*)	48	7.4	1.2 (*S*,*R*)	48	10	1.2 (*S*,*R*)

[b] not determined

To gain first structural insights into the observed regioselectivities, a docking study on selected halohydrin dehalogenases with solved crystal structure was performed. The following crystal structures were chosen as a starting point: i) the (*R*)-para-nitro styrene oxide-bound structure of HheC (PDB: 1ZMT)[185] representing a regioselective enzyme, ii) the substrate-free structure of HheG (PDB: 5O30)[105], as this enzyme was able to convert all five epoxides, and iii) a yet unpublished substrate-free structure of HheD2 representing a non-regioselective enzyme.

De Jong *et al.*[185] could show for HheC that a very specific geometry between the nucleophilic the epoxide ring, the catalytic triad (S132, Y145, and R149) and the nucleophile is required for the nucleophilic attack to occur. This specific orientation determines which of the two carbons of the epoxide ring will react with the nucleophile present in the binding pocket. For HheC, the dockings were in agreement with the experimental data, both substrate enantiomers of **21** and **22** bind in similar productive orientations with the methyl-substituted carbon atom in closer proximity to the nucleophile binding pocket (Figure 22, A and B). HheC, in fact, converted both substrates **21** and **22** with high regioselectivity but poor enantioselectivity and no productive binding was found for substrates **23**, **24** and **25**. In particular, substrates **24** and **25** appear correctly orientated but having the epoxide ring too far away from the catalytic triad. This is again in agreement with the experimental data obtained for HheC, which did not convert epoxides **24** and **25**. Epoxide **23** revealed a more interesting binding mode, in

fact, the terminal hydroxyl group of both substrate enantiomers is pointing towards the catalytic amino acid residues (Figure 22, C). This flipped binding mode of **23** found for HheC might also occur in other HHDHs and could thus explain the lack of activity of most HHDHs toward this epoxide.

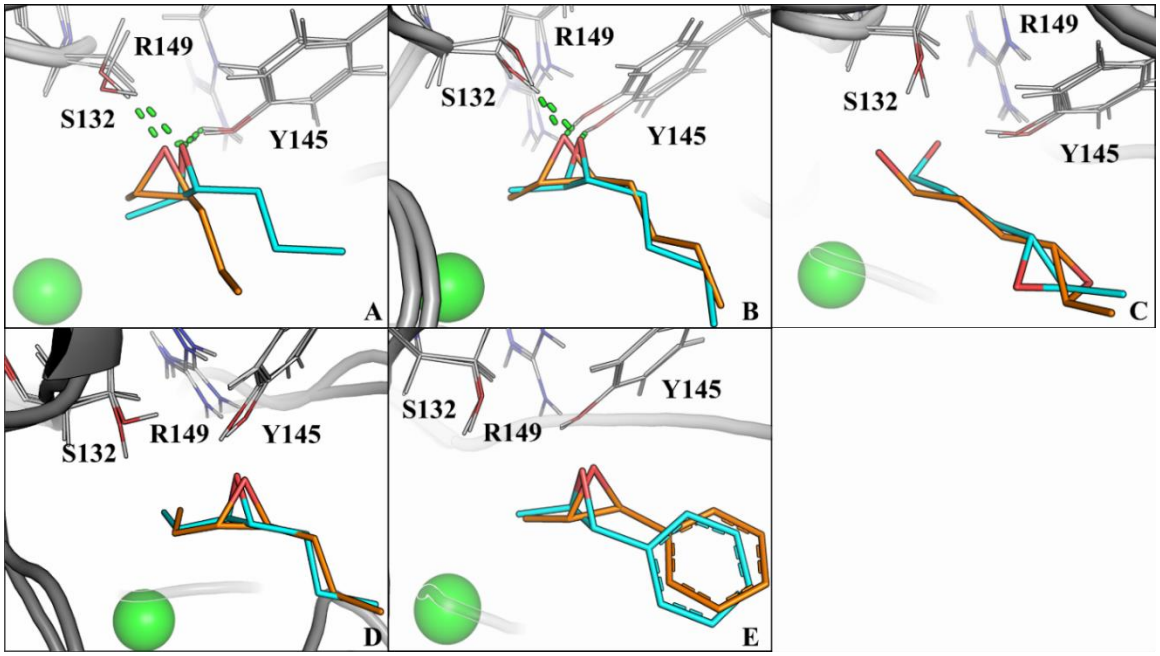


Figure 22. Docking results for HheC (PDB: 1ZMT) with substrates **21** (A), **22** (B), **23** (C), **24** (D) and **25** (E). Substrates with *R,R*-configuration are shown in light blue, whereas substrates with *S,S*-configuration are shown in orange. Hydrogen bonds between epoxide oxygen and catalytic residues S132 and Y145 are represented as green dotted lines. The water molecule present in the nucleophile binding pocket is shown as a green sphere to indicate the position of the nucleophile. Taken from Calderini *et al.*[183]

The same dockings of epoxides **21**-**25** were also performed with HheD2. Docking of substrate **21** resulted in productive binding modes for both enantiomers (Figure 23, A), which is in agreement with experimental data. The binding geometry of (*S,S*)-**21** would favour a nucleophilic attack on the sterically more hindered carbon resulting in the production of (2*S*,3*R*)-**31b**, while the opposite binding pose was not found in the docking study, again in agreement with the experimental data (Figure 23). Epoxide **22** gave productive bindings only for the (*R,R*) enantiomer and interestingly, this-enantiomer showed two possible binding modes with either the longer or shorter substituent entering the active site first. The docking of *trans*-4,5-epoxyhexan-1-ol (**23**) revealed one productive binding mode which is not in agreement with experimental data. Additionally, a flipped binding mode, similar to the one found for HheC (Figure 23, C), was obtained in the docking of (*S,S*)-**23**. This may actually represent the favoured binding mode given that HheD2 displayed no activity in the conversion of this epoxide. Docking of substrate **24** resulted in productive binding modes as opposed to HheC, although both enzymes

showed no or little conversion with this substrate. Once again, different enantiomers have opposite orientations; these poses may occur in other HHDHs from the D phylogenetic subtype and might explain the observed regioselectivity. Finally, both enantiomers of *trans*-1-phenylpropylene oxide (**25**) yielded binding poses that are too far away from the catalytic triad for conversion, which is again in agreement with experimental data.

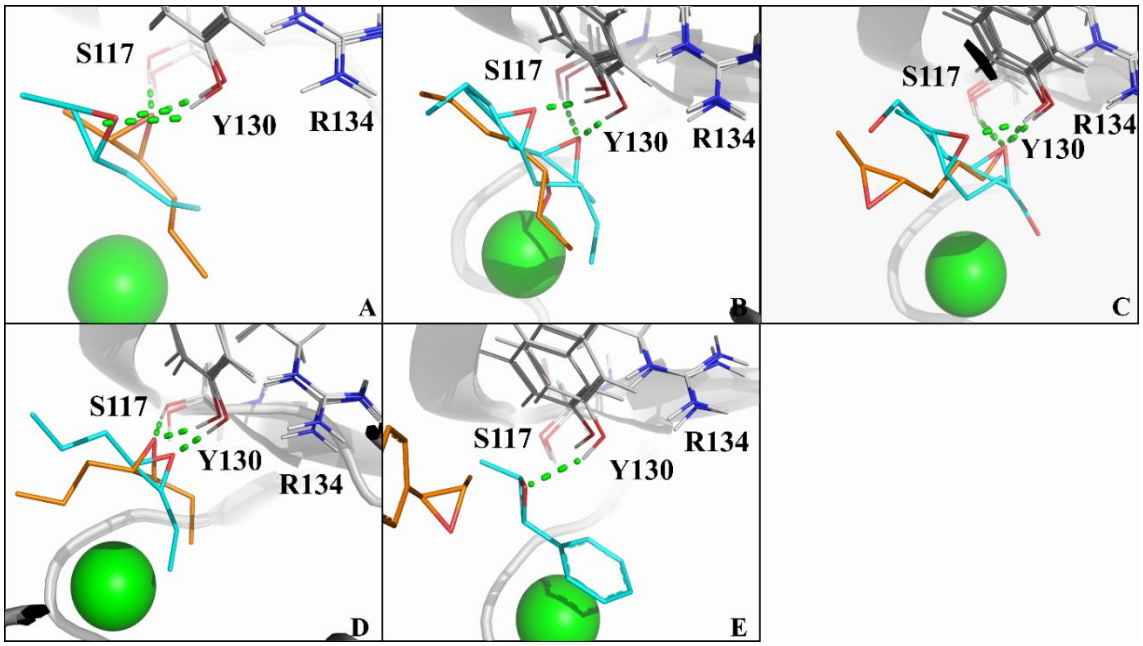


Figure 23. Docking results for HheD2 (unpublished PDB) with substrates **21** (A), **22** (B), **23** (C), **24** (D) and **25** (E). Substrates with *R,R*-configuration are shown in light blue, whereas substrates with *S,S*-configuration are shown in orange. Hydrogen bonds between epoxide oxygen and catalytic residues S117 and Y130 are represented as green dotted lines. The water molecule present in the nucleophile binding pocket is shown as a green sphere to indicate the position of the nucleophile.

The docking results for HheG are overall in agreement with experimental data (Figure 24). However, for epoxides **23**-**25** the docking results look less accurate as not all substrate enantiomers yielded correct geometries. It is important to note that substrates **21**-**24** yielded two possible binding modes, where the longer substituent points in opposite directions. Flipped orientations would favour the nucleophilic attack on opposite carbons of the epoxide ring. This is confirmed by our experimental data as no or only low regioselectivity was observed with these substrates. Dockings of *trans*-1-phenylpropylene oxide (Figure 24, E) resulted in only a productive binding mode for (*S,S*)-**25**. Interestingly, the orientation of both epoxides would favour the nucleophilic attack on the carbon closer to the aromatic ring, in agreement with experimental data. This is also the favoured site of attack in the chemical azidolysis as the intermediate

carbocation is stabilized by electronic resonance with the aromatic ring. Indirectly this may give hints on the regioselectivity observed for HheF. This HHDH may favour the opposite binding mode resulting in a higher ratio of **35a** compared to the other active HHDHs. Although, the overall yield is very low and no crystal structure is available yet to confirm this hypothesis.

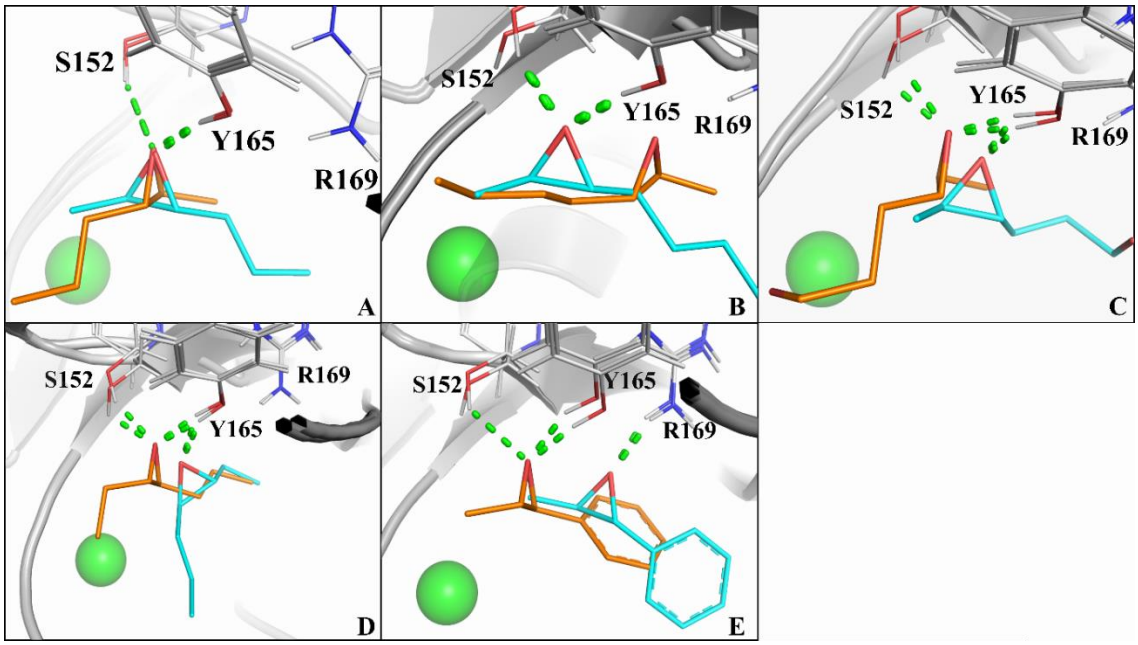


Figure 24. Docking results for HheG (PDB: 5O30) with substrates **21** (A), **22** (B), **23** (C), **24** (D) and **25** (E). Substrates with *R,R*-configuration are shown in light blue, whereas substrates with *S,S*-configuration are shown in orange. Hydrogen bonds between epoxide oxygen and catalytic residues S152 and Y165 are represented as green dotted lines. The water molecule present in the nucleophile binding pocket is shown as a green sphere to indicate the position of the nucleophile. Adapted from Calderini *et al.*[183]

3.1.2.3 Characterization of regioselective halohydrin dehalogenases

After identification of a few regioselective HHDHs in the ring opening of epoxides **21** and **22**, enzymes of subtype E were studied in more detail to determine the most suitable ones for combination with the styrene monooxygenase and/or with the Shi epoxidation diketal catalyst in cascade reactions. For this, the enzymes were applied in the form of whole cells (60 g/L, wet cell weight). Overall, HheE (light blue) and HheE5 (blue) performed evidently better in the conversion of the two substrates. HheE appears to be the most active enzyme in the epoxide ring opening of **21** (Figure 25, A) showing a higher slope in conversion within the first 4 h compared to HheE5. In contrast, HheE5 is slightly faster in the conversion of **22** (Figure 25, B) compared to HheE. Both enzymes are significantly faster in the synthesis of 2-azido-3-hydroxyhexane (**31a**) and 2-azido-3-hydroxyheptane (**32a**) compared to other members of the E-type.

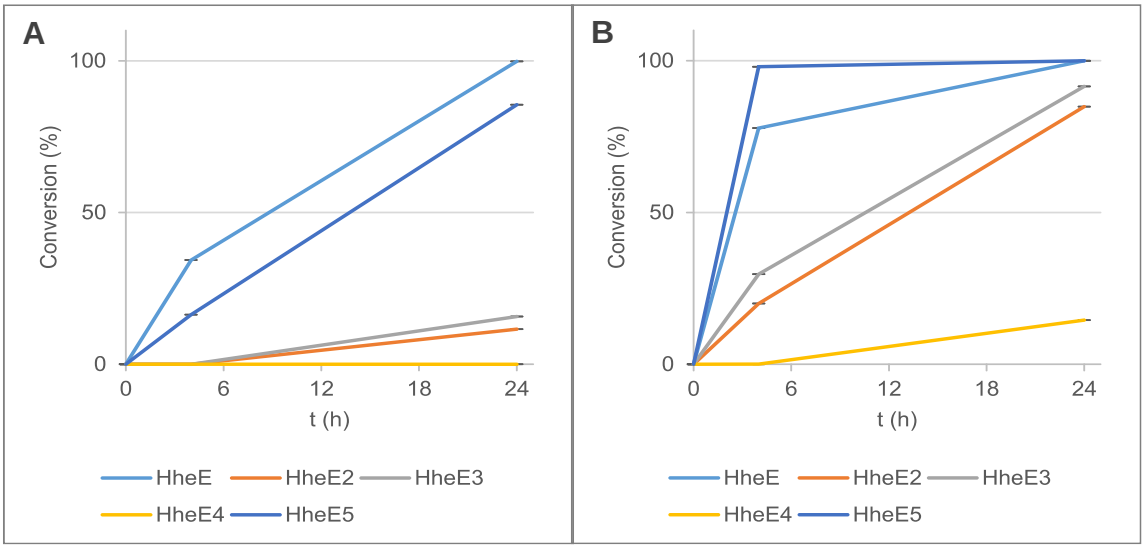


Figure 25. Conversion of epoxide **21** (graph A) and epoxide **22** (graph B) using the regioselective HHDHs from the E-type as whole cells catalysts (60 g/L, wet cell weight). Each reaction was carried out in duplicate using 10 mM of each substrate, 50 mM of NaN_3.

Their kinetic parameters of these two enzymes were determined to further characterise their catalytic properties using epoxides **21** and **22** as well as the nucleophile NaN_3. The reactions were analysed using GC as no reliable colourimetric assay was available. Table 22 shows a summary of the kinetic data obtained using the Hill or Michaelis-Menten equations to fit the experimental data (Appendix, Enzyme kinetics). Both enzymes displayed relatively high K_{50}\K_M values between 18 to 27 mM demonstrating that they possess a rather low affinity for the given substrates. HheE5 exhibited the highest overall maximal reaction velocity as expected from whole cell biotransformations (Figure 25). Both enzymes displayed a higher reaction velocity in the azidolysis of **22** and it is likely that the maximal reaction velocity of HheE5 with epoxide **21** would be lower than the one of HheE giving its lower activity in whole cells reactions. Furthermore, the maximal reaction velocity of HheE in the epoxide ring opening reactions might be influenced by the fixed nucleophile concentration (50 mM) which is ~4-fold higher than its K_M for the nucleophile. In contrast, HheE5's K_{50} for azide is ~8-fold lower than the nucleophile concentration (50 mM) having a better affinity for azide. This may have contributed to the higher V_{max} of HheE5 compared to HheE in the conversion of **22**. The kinetic data of HheE with *trans*-2,3-epoxyhexane (**21**) are lower compared to the ones obtained for **22**; in agreement with whole cells biotransformations as both enzymes converted this substrate less efficiently. Surprisingly no product formation was observed in the first hour and; initial reaction rates were calculated starting from 1h for the following 15 minutes. The Hill coefficient suggests cooperative binding for both substrates (except

for HheE with **21**) which may have been the cause of the observed delay. However, no further analysis was done to elucidate the reasons for this delayed reaction start.

Table 22. Kinetic parameters of HheE and HheE5 for epoxides **21** and **22** using NaN_3 as the nucleophile.

Enzyme	Substrate	K_M /K_{50} (mM)	V_{max} (µmol min^{-1})	kcat (s^{-1})	kcat/ K_M or kcat/K_{50}	n[a]
HheE	**21**	27	55.2	0.75	$2.76 \cdot 10^{-3}$	-[b]
HheE	**22**	18	114.9	0.155	$8.62 \cdot 10^{-3}$	2.3
HheE	NaN_3 (**21**)	12	-	-	-	-[b]
HheE5	**22**	20	152.4	0.206	$10.29 \cdot 10^{-3}$	3.4
HheE5	NaN_3 (**22**)	6	-	-	-	2.6

[a]Hill coefficient ([b] Michaelis-Menten equation was used for the fitting)

3.1.3 Hydrogenation of azidoalcohols

Chemical reduction of an azido group to an amino group is an established reaction that can be performed by a number of metal catalysts.[32,119] Hence, a first selection based on price and reaction conditions was carried out to exclude some of the metals. Several catalysts required extreme conditions, for example the hydrogenation using ruthenium requires 70-80 °C and a pressure of 60-70 atmospheres. These metals were directly excluded as milder conditions were more desirable. In the end, a screening kit containing palladium on different supports was ordered and tested in the conversion of azido-hydroxyheptane (**32**) to amino-hydroxyheptane (**42**). Palladium is a cheap alternative to platinum and it can perform hydrogenations at atmospheric pressure and room temperature, and therefore, it is an ideal catalyst to combine with the previously selected catalysts. In the kit, six different heterogeneous catalysts are provided, which differ in the amount of palladium present and in the material used as support. It is well known that the matrix on which palladium is immobilized can have a great effect on the kind of reaction that is catalysed. A common example is shown in Figure 26, second entry; here the palladium is immobilized on calcium carbonate and it is poisoned with Pb to reduce its activity. This poisoning is typical for catalysts designed to perform hydrogenations on alkynes in order to produce the corresponding alkenes. Therefore, it is not surprising that the lowest conversion was observed using this catalyst. Overall, the fastest heterogeneous catalyst to synthesize amino-hydroxyheptane (**42**) was palladium on carbon (entry 1 and 5 Figure 26). Palladium hydroxide on carbon (Figure 26, entry 3) appeared to have a good reaction rate too. However, it contained twice as much palladium (20 wt. %) compared to the catalyst in entry 1 or 5. Finally, Palladium (10 wt. %, entry 1 in Figure 26) on activated carbon was chosen as the catalyst for combination with steps one and two of the cascade.

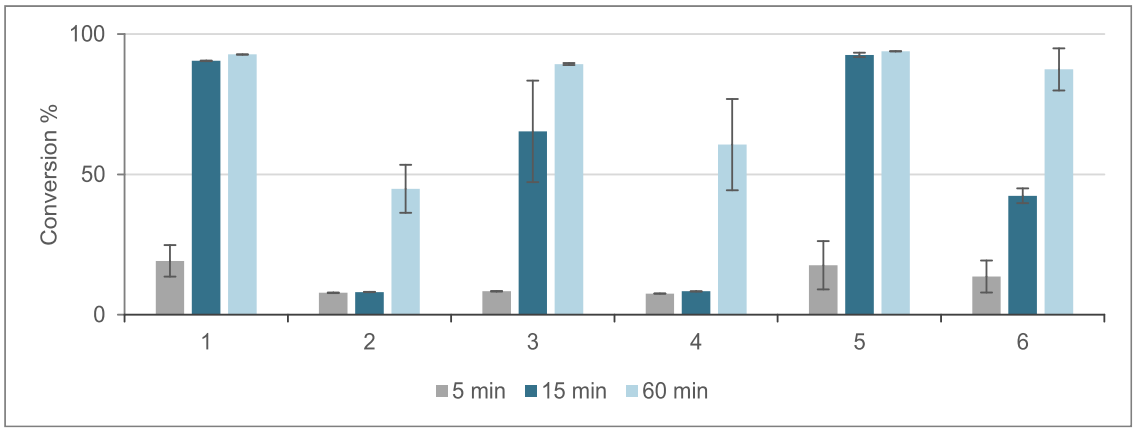


Figure 26. Conversions of the racemic azido hydroxyheptane (**32**) by Pd on different supports. 1 Palladium on carbon, 10 wt. % loading, matrix activated carbon support. 2 Palladium on calcium carbonate, 5 wt. % loading, poisoned with lead. 3 Palladium hydroxide on carbon, 20 wt. % loading (dry basis), matrix carbon, wet support. 4 Palladium on barium sulphate, reduced, 5 wt. % loading. 5 Palladium on carbon, 10 wt. % loading (dry basis), matrix activated carbon, wet support, Degussa type E101 NE/W. 6 Palladium on alumina, 10 wt. % loading, powder, reduced, dry.

An interesting alternative option for the hydrogenation step is to immobilize palladium nanoparticles inside apoferritin. Ferritin is a protein known to store iron in several bacteria. Hence, it is possible to exploit this function to encapsulate different metals including palladium inside the ferritin cage. This would maximize the compatibility of the metal catalyst with the enzyme catalysts of the cascade as the enzymes would be separated from the metal and *vice versa*. The synthesis of this catalyst requires several steps (section 2.2.4, Hydrogenation): at first the apoferritin is incubated with palladium salt to promote the encapsulation. Afterwards, the excess of palladium is washed using a PD10 column, and lastly the mixture is incubated under an H_2 atmosphere for a few hours. The picture in Figure 28, shows the last step of the synthesis. In particular, hydrogen is bubbled through the solution of palladium entrapped into the apoferritin cage. This step is crucial as the reduction of the metal leads to aggregation and formation of the nanoparticles. The aggregation is reflected in the change of colour from pale yellow to dark grey. To further confirm the successful palladium encapsulation, a microscopic characterization was kindly carried out by Dr. Wuyuan Zhang (TU Delft). Figure 27 shows a TEM picture of the nanoparticles. Particularly on the left picture, it is possible to see the homogeneous size distribution of the particles calculated to be 5 ± 1 nm in size. Additionally, the higher magnification picture on the right allows to see apoferritin from horse spleen (AFHS) in white and the palladium nanoparticles as dark spots in inside the AFHS.

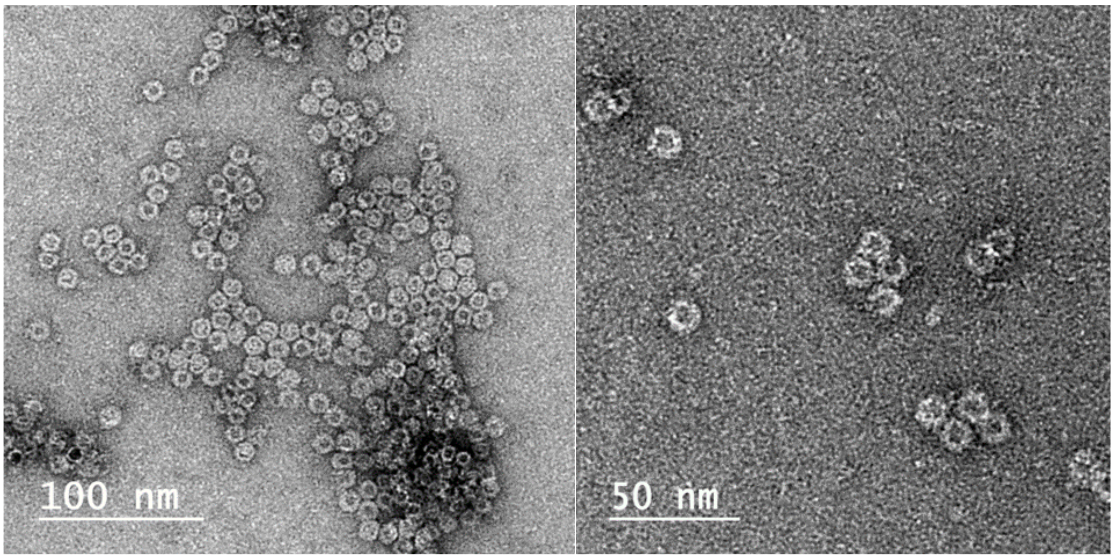


Figure 27. TEM images of Pd@AFHS nanoparticles stained with uranyl acetate at different scale (100 nm and 50 nm). The nanoparticles show a very narrow size distribution of 5 ± 1 nm. The picture on the right clearly shows the nanoparticles with a white coating that represents the apoferritin from horse spleen (AFHS) and a black core composed of Pd.

The graph in Figure 28 shows the conversion of 5 mM racemic 2-azido-3-hydroxyheptane (**32**) to 2-amino-3-hydroxyheptane (**42**) over time. After 4 h, the reaction was almost complete using an amount of Pd@AFHS corresponding to 3.5 µM of Pd. Despite this successful conversion, Pd@AFHS nanoparticles were not considered for further application as the synthesis of this catalyst is quite time consuming. Additionally, apoferritin from horse spleen is an expensive component. Therefore, palladium on carbon was chosen for further experiments having in mind a possible commercial exploitation of the cascade in the future.

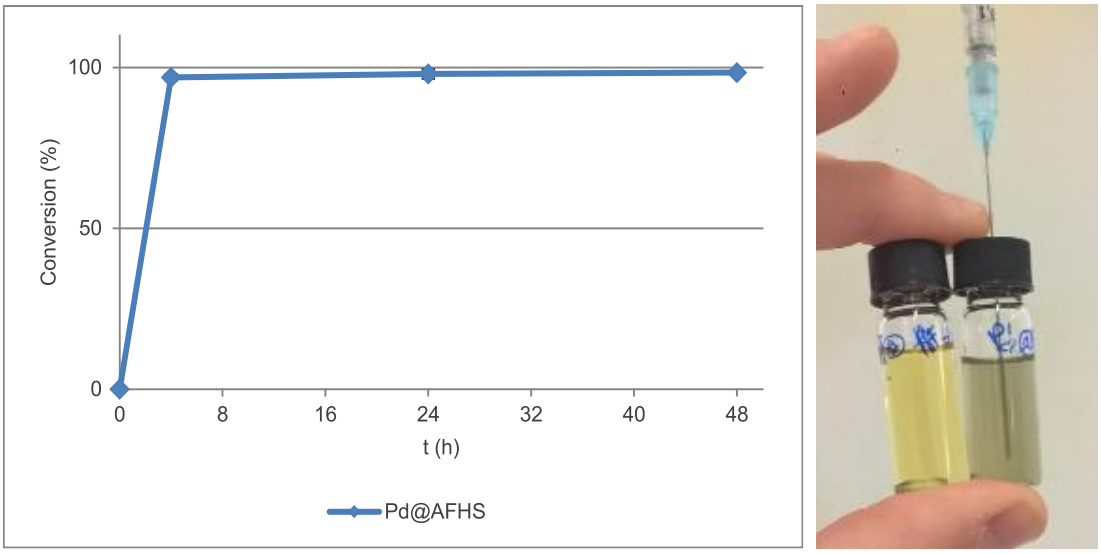
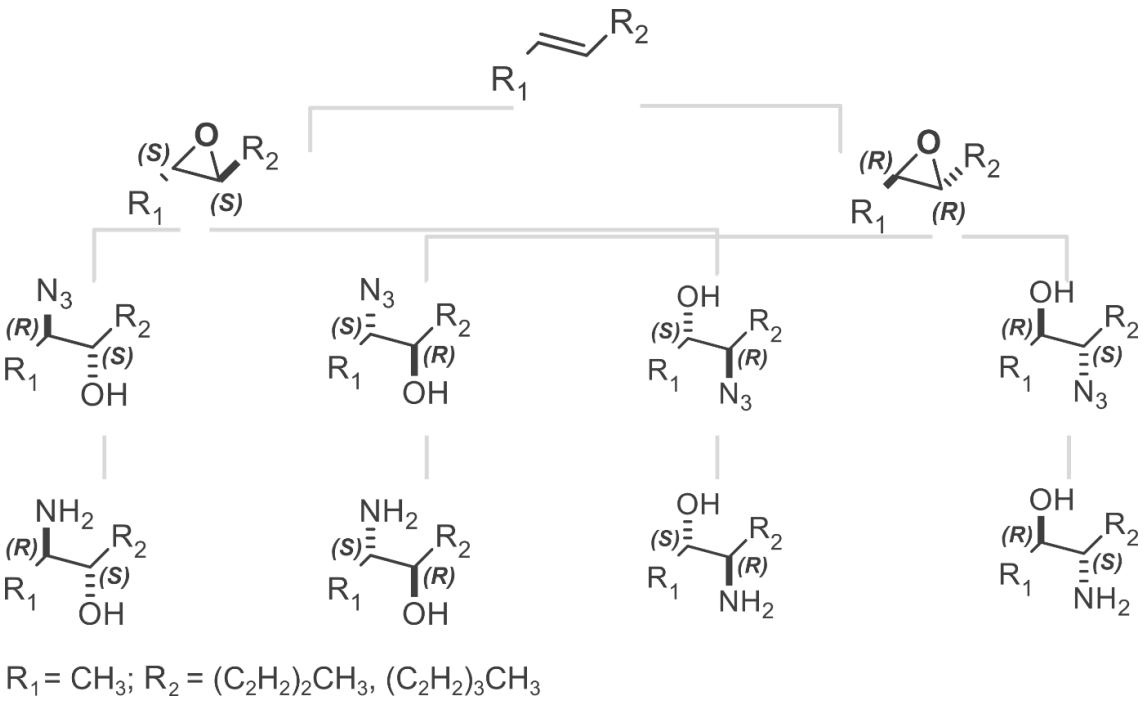


Figure 28. Conversion of 5 mM racemic azido-hydroxyheptane (**32**) to amino-hydroxyheptane (**42**) using Pd nanoparticles entrapped in apoferritin from horse spleen (AFHS). On the right, the photo shows the solution

of Pd and AFHS before and after the reduction of Pd using H_2 that causes the aggregation of Pd in nanoparticles that causes the solution to turn dark grey.

3.1.4 Catalyst selectivity and assignment of azido alcohol stereoisomers



Scheme 19. The epoxidation of a non-terminal *trans*-alkene generates two new stereocenters. Consequently, in the epoxide ring opening, following the typical S_N2 mechanism, four possible azido alcohols are obtained. Lastly, hydrogenation of the azido alcohols does not affect the stereocenters, hence, yielding the four corresponding amino alcohols.

Scheme 19 illustrates the formation of all possible product isomers of a fully unselective cascade for amino alcohol synthesis starting from non-terminal alkenes. Epoxidation of a non-terminal *trans*-alkene will yield the corresponding epoxide in (*S*,*S*)- and (*R*,*R*)-configuration. Subsequently, unselective epoxide ring opening following an S_N2 mechanism will result in the formation of four different azido alcohol stereoisomers: (2*R*,3*S*)- and (2*S*,3*R*)-2-azido-3-alcohol as well as (2*S*,3*R*)- and (2*R*,3*S*)-3-azido-2-alcohol. Finally, hydrogenation of the azido groups will not affect the stereochemistry of the compounds. According to our strategy, having a stereoselective epoxidation would reduce the possible azido- and aminoalcohols isomers from four to two. Consequently, a regioselective epoxide-ring opening would further reduce the possibilities from two to one, thus, leading to an enantiopure product.

To confirm this, a first cascade was set up involving the enantioselective styrene monooxygenase and a non-regioselective HHDH. In particular, for *trans*-2-hexene (**1**) as

substrate StyAB was coupled with HheG and for *trans*-2-heptene (**2**) StyAB was coupled with HheD5. As shown in section 3.1.1.4, StyAB only produces (*S*,*S*)-*trans*-2,3-epoxyhexane (**21**) or (*S*,*S*)-*trans*-2,3-epoxyheptane (**22**). Thus, coupling it with an unselective HHDH should result in the production of similar quantities of (2*R*,3*S*)-2-azido-3-hydroxyhexane (**31a**) and (2*S*,3*R*)-3-azido-2-hydroxyhexane (**31b**) or (2*R*,3*S*)-2-azido-3-hydroxyheptane (**32a**) and (2*S*,3*R*)-3-azido-2-hydroxyheptane (**32b**). This is explained by the catalytic mechanism of HHDHs, which leads to the inversion of the stereocenters due to nucleophilic attack following an S_N2 reaction. Hence, epoxide ring opening of the (*S,S*)-enantiomer produced by the styrene monooxygenase leads to an inversion of the attacked stereocenter yielding the corresponding azido alcohol in (2*R*,3*S*)- or (2*S*,3*R*)-configuration for regioisomer **a** and **b**, respectively. Comparison of the chiral GC data from these cascade reactions with the regioselectivity data obtained from GC-MS in the ring-opening experiments (section 3.1.2.2) would allow the assignment of the corresponding stereoconfiguration of the four peaks.

The respective chiral GC chromatograms for the two cascade reactions combining StyAB with HheG (A) or HheD5 (B) are shown in Figure 29. These reactions were run in a sequential mode, the epoxidation for the first 8 h, and afterwards the respective HHDHs and sodium azide were added and the epoxide-ring opening was run for further 16 h. Interestingly, both enzymes displayed some regioselectivity in this cascade reaction. Originally, the two HHDHs were selected as they formed both product regioisomers in similar amounts (regioisomeric ratio of 50:50) in the conversion of racemic epoxides **21** and **22**, respectively (Appendix, Supporting table 1). However, when starting from the corresponding (*S,S*)-epoxides, the determined regioisomeric ratio was 40:60 (**31a**:**31b**) and 22:78 (**32a**:**32b**) for HheG and HheD5, respectively. This means that both enzymes preferentially catalysed the nucleophilic attack on the sterically more hindered carbon atom leading to higher quantities of the (2*S*,3*R*)-enantiomer.

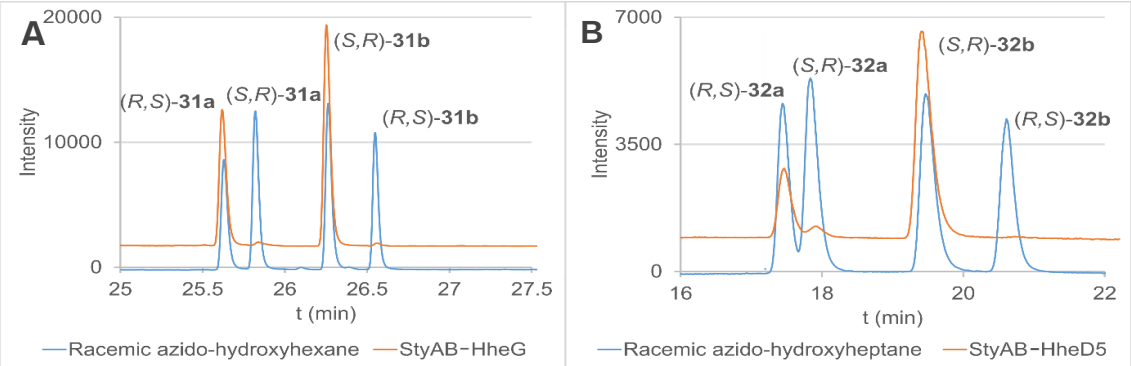


Figure 29. Comparison of chiral chromatograms of the azidoalcohols chemically produced starting from racemic epoxides (blue) or synthesised in starting from (S,S)-**21** (A) or (S,S)-**22** (B) synthesised by StyAB (orange). The epoxide-ring opening was catalysed by HheG for the production of (2*R*,3*S*)-2-azido-3-hydroxyhexane (**31a**) and of (2*S*,3*R*)-3-azido-2-hydroxyhexane (**31b**) (orange line in A) and by HheD5 for

the production of (2*R*,3*S*)-2-azido-3-hydroxyheptane (**32a**) and of (2*S*,3*R*)-3-azido-2-hydroxyheptane (**32b**) (orange line in B).

To validate the peak assignment, cascades reactions coupling enantioselective epoxidation with regioselective epoxide-ring opening were carried out. This way, the reactions would afford a single diastereoisomer with the azido group in the sterically less hindered carbon (e.g. **31a** and **32a**). In particular, coupling the Shi diketal epoxidation with either HheE or HheE5 (Figure 30, orange line) resulted in higher amounts of (2*S*,3*R*)-2-azido-3-hydroxyhexane (**31a**), (2*S*,3*R*)-2-azido-3-hydroxyheptane (**32a**) respectively. In this case, both enantiomers are visible as the enantiomeric excess values obtained were 60% and 48%, respectively. In contrast, coupling the styrene monooxygenase with the same HHDHs (Figure 30, grey line) resulted in the production of the opposite enantiomers, (2*R*,3*S*)-2-azido-3-hydroxyhexane (**31a**), (2*R*,3*S*)-2-azido-3-hydroxyheptane (**32a**), respectively. The styrene monooxygenase showed significantly higher enantioselectivity, hence only one peak is visible. Furthermore, the peak order is reversed compared to Figure 29, this is due to the different columns used for the chiral separation. The chromatograms in Figure 29 have been generated at TU Braunschweig using a Lipodex E chiral column while the chromatograms in Figure 30 have been generated at Enzymicals, Greifswald using a Lipodex G column. The inversion is very likely due to the different matrix compositions.

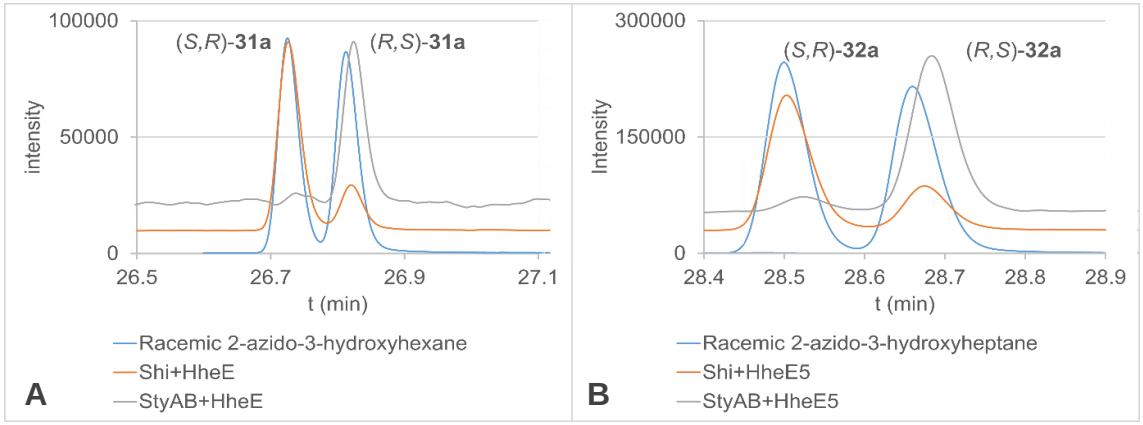
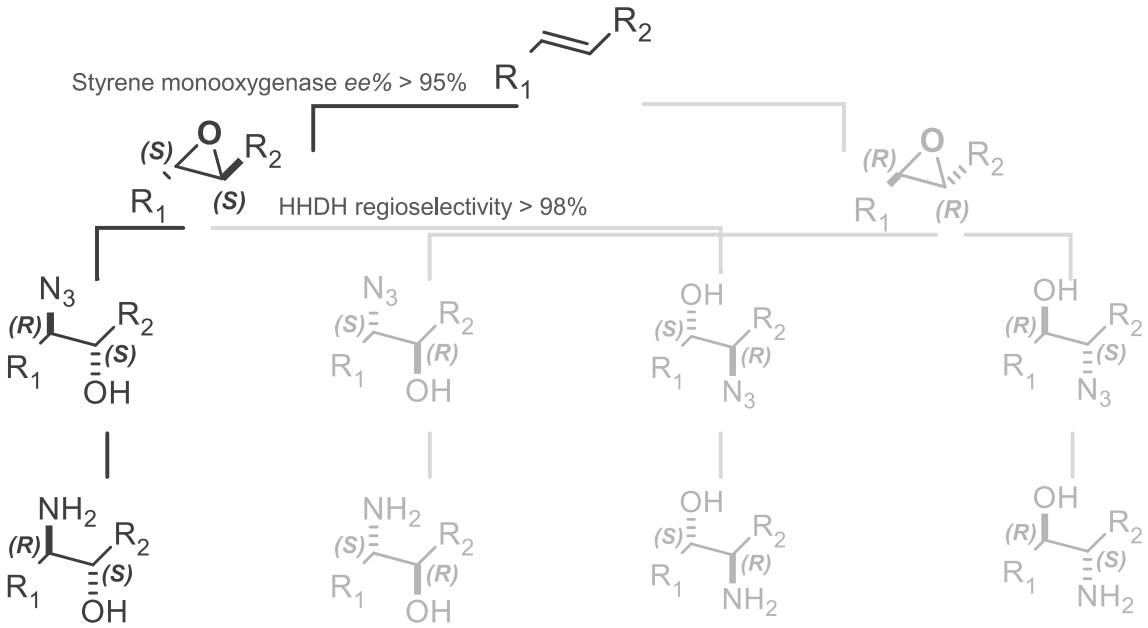


Figure 30. Chiral chromatograms of 2-azido-3-hydroxyhexane (A) and 2-azido-3-hydroxyheptane (B). The grey line represents **31a** (A) and **32a** (B) produced in cascades composed by the styrene monooxygenase and HheE (A) or HheE5 (B). These reactions preferentially produce (2*R*,3*S*)-**31a** (A) or (2*R*,3*S*)-**32a** (B). The orange line represents **31a** (A) and **32a** (B) produced in cascades composed by the Shi epoxidation diketal catalyst and HheE (A) or HheE5 (B). These reactions preferentially produce (2*S*,3*R*)-**31a** (A) or (2*S*,3*R*)-**32a** (B). The blue line represents racemic azidoalcohols **31a** (A) and **32a** (B) for reference.

3.1.5 StyAB – HHDH – Pd/C cascade

According to our strategy, coupling a stereoselective epoxidation step with a regioselective epoxide ring opening step will yield an enantiopure azido alcohol. Lastly, reduction of the azido group produces the corresponding amino alcohol. In the first cascade, the styrene monooxygenase, synthesising epoxides (*S,S*)-**21** and (*S,S*)-**22,** is combined with regioselective HHDHs HheE (for **21**) or HheE5 (for **22**) and Pd/C as hydrogenation catalyst. This catalyst combination allows the selective synthesis of respective azido- and aminoalcohols in (2*R*,3*S*)-configuration (Scheme 20).

Each reaction, composing the cascade, was first optimized separately and cross-interactions were studied before setting up the cascade. Cross-interaction studies are important to determine components that could have a negative effect on other catalysts of the cascade. This allows finding an effective compromise to maintain reasonable working conditions for all catalysts.



Scheme 20. General scheme of the cascade composed of the styrene monooxygenase in the first step, a halohydrin dehalogenase in the second step and palladium on carbon in the last step. Enantioselective epoxidation of the *trans*-alkene combined with regioselective epoxide ring-opening leads to the formation of only one out of four possible stereoisomers enabling the synthesis of an enantiopure amino alcohol.

3.1.5.1 Optimization and scale-up of the StyAB-catalysed epoxidation

At first, a design of experiment (DOE) optimization was performed in order to identify critical variables of this reaction step. Design of experiment enables the identification of important variables as well as determining optimal combinations by statistical analysis

(DOE plugin in the software for statistical analysis R). Four variables were selected for the epoxidation using whole cells and glucose as co-substrate for cofactor regeneration: cellular concentration, glucose concentration, stirring speed and temperature. For each of the variables two values were examined and the product concentration after two hours was used as response value. The main effects plot is given in Figure 31 showing that for two variables (OD_{600} and rpm) significantly steeper slopes are obtained compared to the rest. Of these, the cellular concentration had the strongest negative effect on the product concentration in agreement with previous results (Figure 18). In contrast, higher stirring speed showed a positive effect on the product concentration with a 9% improvement. The glucose concentration exhibited a negative slope as seen for the cellular concentration indicating that adding more glucose for cofactor regeneration probably boosts the cellular metabolism and as a result, the oxygen availability for the enzyme is reduced. Overall, most of these effects can be explained with oxygen limitation. A higher concentration of biomass consumes more oxygen, which is also required by the monooxygenase (see below). A faster stirring speed can improve the oxygen resident time in solution and at the same time, it can reduce mass transfer limitations. Hence, substrate and oxygen are better distributed in the reaction mix and higher conversions were reached. Finally, the temperature had only a minor effect, but slightly less epoxide product was formed at 30 °C. This could point once more at an insufficient oxygen availability as gasses are normally more soluble in aqueous systems at lower temperatures. The obtained conversions, however, are very low compared to the analytical scale reactions where almost full conversion was achieved already after 1 h. Here, as visible in Figure 31, a maximum product concentration of 0.185 mM was reached after 2 h, which corresponds to a conversion of 3.7%. Very likely, mass transfer limitation together with oxygen availability appeared are the important parameters to optimize order to improve epoxide production in the StyAB-catalysed step.

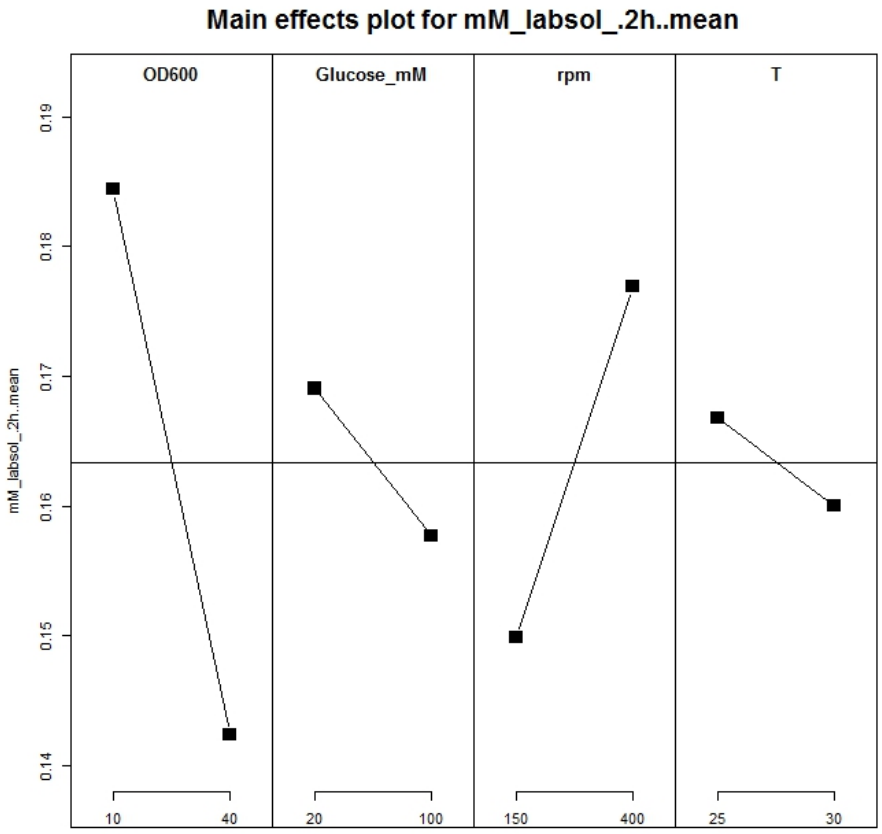


Figure 31. Main effects plot for the DOE for optimization of the StyAB-catalysed epoxidation of *trans*-2-heptene (5 mM). Two values for each of the four variables were taken into account. The concentration of the epoxide after 2 h was selected as response factor. All reactions were performed in duplicate.

No further progress in the epoxide production could be made by only relying on cofactor regeneration by cellular metabolism. Hence, an enzymatic cofactor regeneration catalysed by ADH was tested in combination with SyAB. As mentioned in section 3.1.1.2, the alcohol dehydrogenases from *Lactobacillus kefir* (LkADH) and *Leifsonia sp.* S749 (LsADH) were selected as cofactor regeneration enzymes. At first, StyAB and LkADH were produced separately and mixed together for the reaction. Using cell-free extract and cofactor regeneration could already improve the epoxide production in comparison with whole cells reactions where cellular metabolism was used for cofactor regeneration (Figure 32, WC - 40mL and first entry for CFE - 80 mL). Additional improvements were achieved by using substrate feed (Figure 32, Orange No S feed vs S feed) confirming that StyAB suffers from substrate inhibition. Unfortunately, LkADH, when coexpressed, almost completely dominated enzyme production resulting in very low StyAB levels (Section 3.1.1.2) and no product formation (data not shown). In contrast, LsADH, showing a more balanced expression, could reach a product formation comparable with experiments where LkADH was added separately (Figure 32, LkADH / S feed vs LsADH coexpressed / S feed). In these experiments, a lower overall conversion was reached

compared to LkADH added separately, but this way the total protein content could be reduced as two different cell-free extracts are not mixed.

In small scale experiments, it appeared that StyAB's activity drops to almost zero in 6 h (Figure 19, Section 3.1.1.5), hence enzyme feed was tested. Feeding cell-free extract as well as substrate could further improve epoxide production (Figure 32, Red S feed vs Enzyme + S). At last as done previously,[68,84] an organic phase acting as substrate and product sink was tested to see if it would improve the final epoxide concentration. The presence of a second organic phase resulted in a 4-fold improvement compared to the second highest product concentration reached (Figure 32, LsADH coexpressed / Enzyme + S/ 25% organic phase vs LkADH / S feed). In conclusion, using cell-free extract and enzymatic cofactor regeneration could improve the epoxide production compared to whole cells experiments with metabolism as cofactor regeneration. Enzyme and substrates feed further improved production, however the presence of an organic phase showed the highest positive effect among all variables tested (Figure 32, blue).

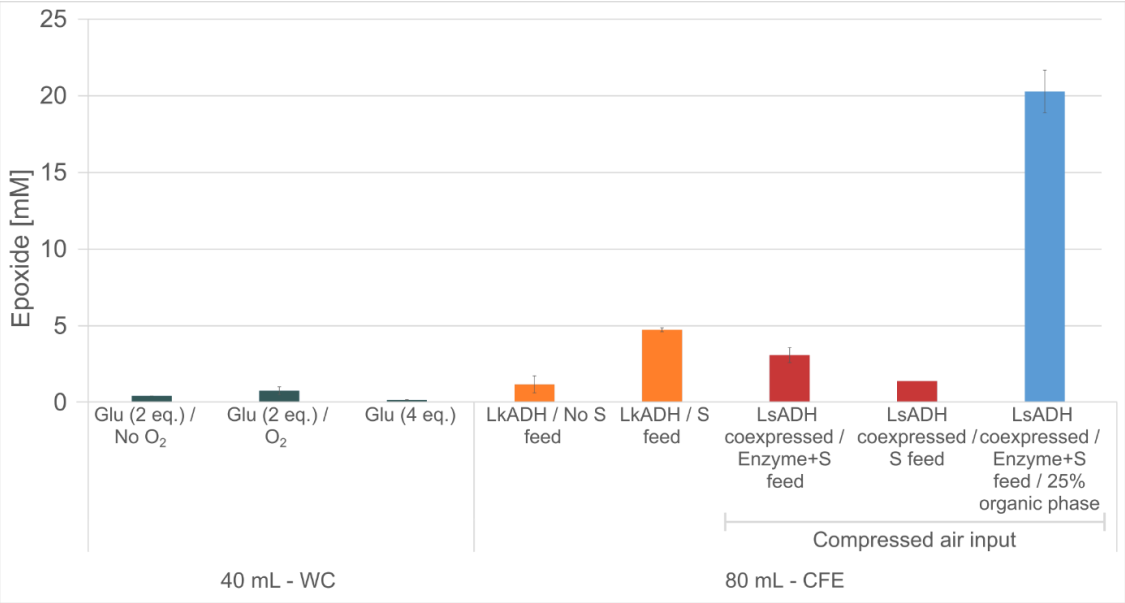


Figure 32. Different experimental conditions used for the epoxidation of *trans*-2-heptene (**2**) catalysed by StyAB. Reactions using StyAB as CFE and LkADH for cofactor regeneration added as CFE separately with or without substrate feed are given in (Orange). Reactions using StyAB coexpressed with LsADH as CFE with substrate feed and with or without enzyme feed are given in red. The same reaction (with enzyme feed) using 25% heptane as organic phase is shown in blue.

The best conditions among the tested setup were achieved by using enzymatic cofactor regeneration, enzyme and substrate feed, an organic phase and compressed air input. The beneficial effect of compressed air could not be quantified with these experiments, however the presence of oxygen appeared to be an extremely important variable to control during the reaction. Thus, an oxygen sensor was used to measure O_2 consumption in reactions with cell-free extract and enzymatic cofactor regeneration or whole cells and cellular metabolism as cofactor regeneration to further investigate the

oxygen limitation. Figure 33 A shows that the O_2 concentration dropped to 0 µmol/L in about 7 minutes when whole cells (OD_{600}=10, corresponding to 15 g/L wet cell weight) were used as biocatalyst. In contrast, in reactions containing cell-free extract (from OD_{600}=10, corresponding to 2.3 g/L total protein) the oxygen concentration dropped to around 80 µmol/L without compressed air input. This could explain the extremely low epoxide formation reached in whole cells reactions, as cellular metabolism would consume all the oxygen. Adding compressed air (3 L/h) maintained the oxygen saturated in cell-free extract reactions for the entire tested time, whereas whole cells consumed oxygen at a faster rate (than 3 L/h) keeping the oxygen concentration below 100 µmol/L for the first seven hours (Figure 33, B). However, the difference in oxygen consumption between the two reactions does not correlate with the final product concentrations, as only 0.9 mM epoxide were formed in the reaction containing whole cells and 1.4 mM in the CFE reaction. In this case, a second organic phase could not be used as the oxygen sensor would be damaged by organic solvents.

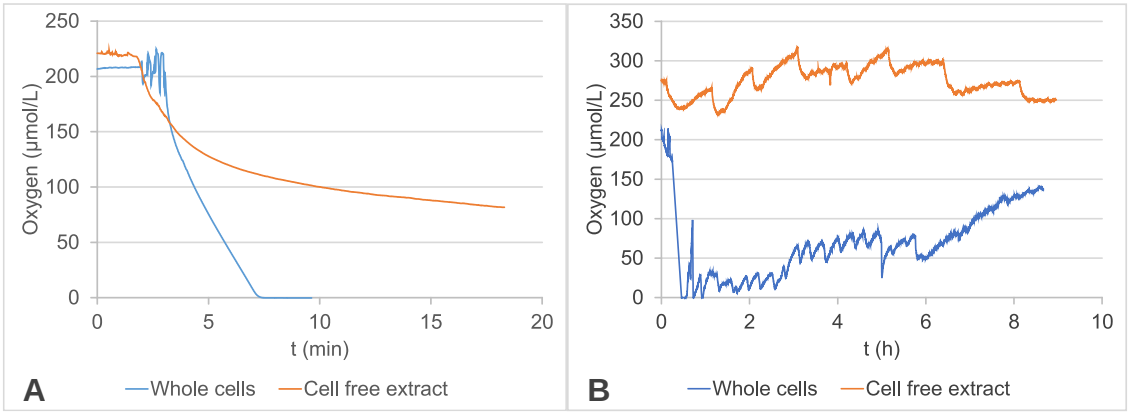


Figure 33. Oxygen consumption in styrene monooxygenase-catalysed reactions using whole cells and cofactor regeneration by cell metabolism (orange) or cell-free extract harbouring StyAB and coexpressed LsADH for cofactor regeneration (blue). The curves show the oxygen concentration in solution after addition of the respective co-substrates for cofactor regeneration (glucose and isopropanol, respectively). In **A** no compressed air was added, whereas in **B** compressed air was constantly introduced at a flow of 3 L/h and the oxygen concentration was measured for longer times.

The final epoxide concentration does not reflect the final epoxide concentration reached by the two reactions system, hence apparent kinetic data using different O_2 concentrations were measured for StyAB co-expressed with LsADH applied as cell-free extract.

The O_2 concentration was adjusted by removing excess O_2 by bubbling nitrogen in the solution until a user-defined O_2 concentration (7,10, 20, 30, 40, 100, 150 µmol/L). Afterwards, the time needed for the oxygen to diffuse back into the solution is measured after addition of the substrate. This time is then compared with the diffusion rate of oxygen when no substrate is added. The differential rate between these two

measurements is the enzymatic reaction; as the monooxygenase reaction would consume some of the oxygen diffusing in solution, hence O_2 would saturate the solution faster when no substrate is present.[186] StyAB displayed a good affinity for O_2 with a rather low K_M of 14.6 µmol/L. As a rule of thumb oxygen concentration should be kept above ten times the K_M (approx. 146 µmol/L) to work at velocities close to V_{max}. This could explain the lower productivity of StyAB when whole cells are used as the biocatalyst.

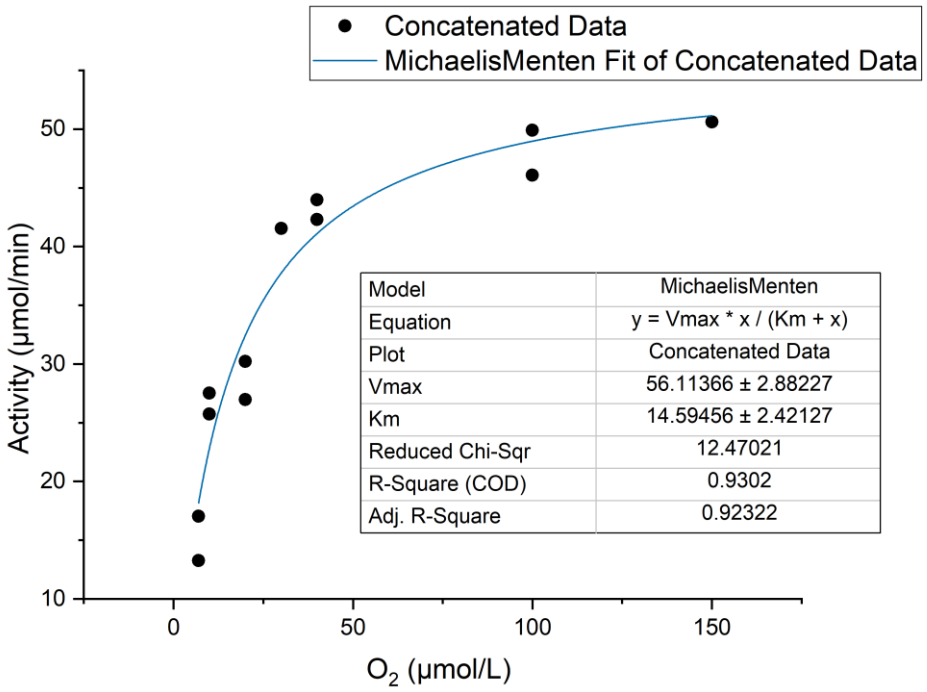


Figure 34. Apparent K_M and V_{max} measured for StyAB coexpressed with LsADH applied as cell-free extract using different O_2 concentrations. Kinetic data were collected in duplicate.

Figure 32 (blue) demonstrated that the use of a second organic phase (25% v/v) could improve *trans*-2,3-epoxyheptane (**22**) production 4-fold compared to the second highest product concentration reached (Figure 32, LsADH coexpressed / Enzyme + S/ 25% organic phase vs LkADH / S feed). Therefore, different organic solvents were tested in order to find the best for the epoxidation of *trans*-2-heptene (**2**) as a model substrate. Several organic solvents ranging from hydrophobic (isooctane) to polar (ethyl acetate) were tested using the optimal setup found (StyAB coexpressed with LsADH, enzyme and substrate feed and compressed air input and 25% organic phase). Overall the epoxidation of **2** seems to be improved by a hydrophobic second phase (Figure 35), while the polar solvents tested did not yield any significant amount of epoxide product

(data not shown). Among them, pentyl acetate and hexyl acetate are worth mentioning as StyAB was reported to reach high product concentrations on small scale in presence of these solvents.[68] On big scale, however, all polar solvents caused considerable protein denaturation in short time leading to extremely low product concentrations (data not shown). Adding 100 µmol NADH and 50 µmol FAD did not seem to affect the final product concentration. In conclusion, heptane appeared to be the best compromise as it would allow better downstream processing having a lower boiling point compared to isooctane and would be easier to remove in vacuum. Finally, addition of expensive cofactors can be omitted since they did not show any positive effect on the final product concentration.

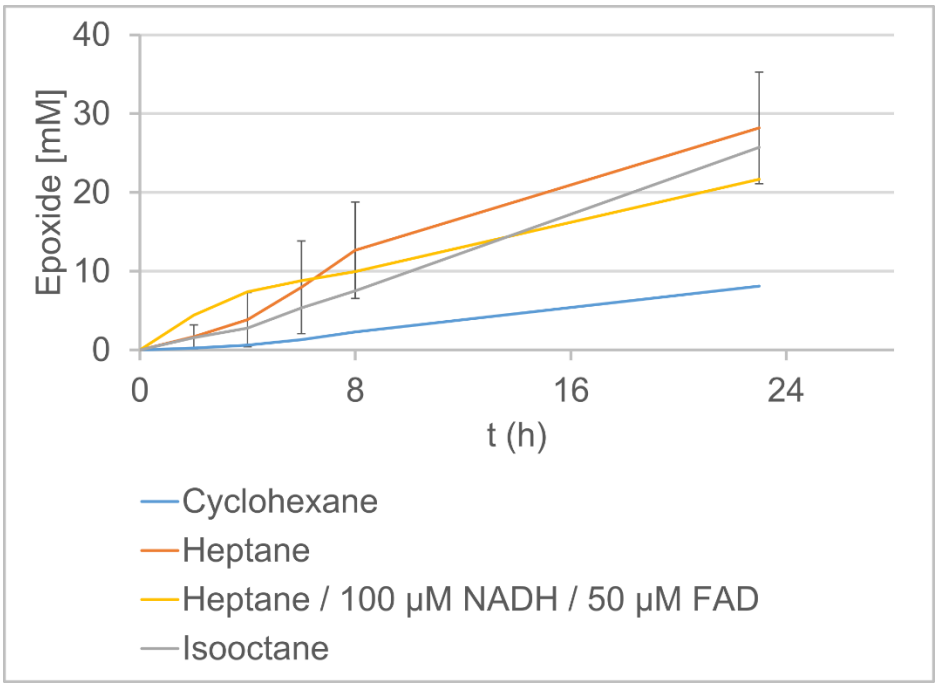


Figure 35. Comparison of different organic solvents used as second phase for the production of *trans*-2,3-epoxyheptane (**22**). All reactions were performed using the best setup found (StyAB coexpressed with LsADH, enzyme and substrate feed and compressed air input and 25% organic phase).

Heptane could successfully act as a substrate and product reservoir for the styrene monooxygenase as it would be necessary to reach a total epoxide concentration of 100 mM to reach 20 mM in the aqueous phase (Figure 36). On the other hand, heptane might be a problem for the halohydrin dehalogenase as a low epoxide concentration would be present in the aqueous phase (under 10 mM) giving that the styrene monooxygenase on average could reach a product concentration of 28 mM. Interestingly, it appeared that epoxide **22** concentration in the aqueous phase is higher than **21**, although epoxide **22** as a longer chain that should make it more hydrophobic, whereas *trans*-2-hexene (**1**) is the least soluble in water compared to the epoxides. The partition of *trans*-2-heptene (**2**)

could not be determined due to analytics limitation, as heptane has a similar retention time, in GC, hence it covers *trans*-2-heptene (**2**) peak completely.

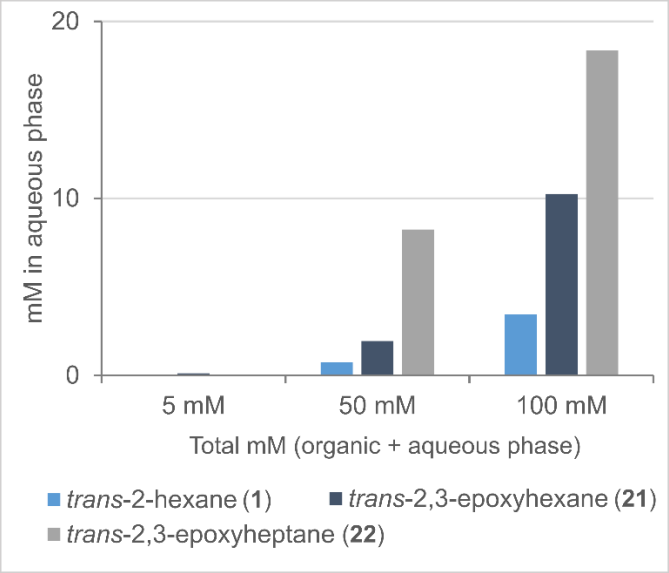


Figure 36. The graph shows the concentrations of the different compounds found in the aqueous phase when 50% (v/v) heptane was used. When the compounds were present at a concentration of 5, 50 or 100 mM in aqueous phase before extraction.

After optimising the reaction system for epoxidation of *trans*-2-heptene, the optimised setup (Figure 32, blue) was also tested for the synthesis of *trans*-2,3-epoxyhexane (**21**) in 80 mL scale and in larger-scale reactions (250 ml) for the conversion of *trans*-2-heptene (**2**). The reactions were carried out at 25 °C and 300 rpm, with heptane as second phase (25% v/v). Substrate and enzyme were fed hourly until a final concentration of 100 mM and 9.2 g/L (total protein concentration), respectively. In Figure 37, the orange curve shows the epoxide formation in a reaction using *trans*-2-hexene (**1**) as substrate, where only a low epoxide concentration (6 mM) could be achieved compared to the epoxidation of *trans*-2-heptene under similar conditions (Figure 35).

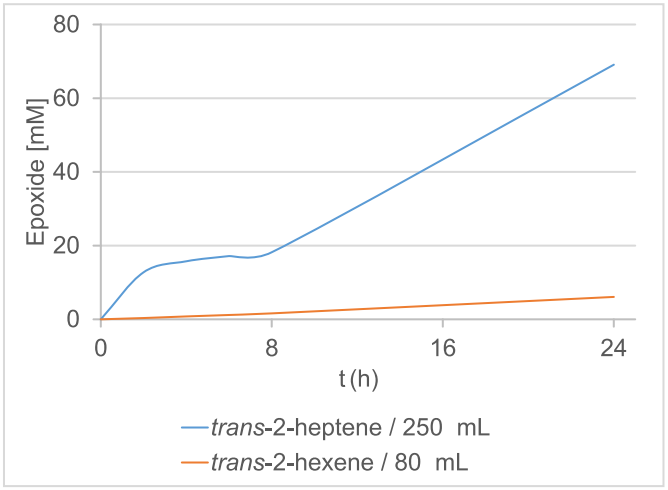


Figure 37. StyAB coexpressed with LsADH, enzyme and substrate feed and compressed air input and 25% organic phase were used for the production of *trans*-2,3-epoxyhexane (**21)** (Orange). The blue line shows the same reaction setup for the epoxidation of *trans*-2-heptene (**2**) on 200 mL scale (blue line).

On the other hand, the blue curve (Figure 37) shows the result of *trans*-2-heptene (**2**) epoxidation on 200 mL scale, for which surprisingly a final epoxide concentration of 69 mM (69% conversion) could be achieved. This represents the highest product concentration reached in this work, which, however, could not be reproduced in further tests. This might be due to varying expression levels of StyAB in the CFE used as catalyst. All in all, the relatively low StyAB productivity would very likely limit the overall final product concentration when used in cascade reactions. Nonetheless, the reaction conditions could be significantly improved to reach acceptable epoxide concentrations for a proof of concept of the whole cascade for enantiopure amino alcohol synthesis.

3.1.5.2 Scale-up of the epoxide-ring opening step

All HHDH experiments were performed using 5 mM substrate concentration, however higher substrate concentrations are desirable during scale up. Whole cells (60 g/L wet cell weight) harbouring either HheE or Hhe5 were used to convert racemic epoxides **21** and **22**, respectively, ranging from 50 to 300 mM. In all cases, two equivalents of the nucleophile (NaN_3) were added (Figure 38). HheE showed reasonable conversions compared to the chemical background (dashed lines) only using 50 mM substrate (**21**) concentration (Figure 38, A). Using 200 and 300 mM of epoxide **21** resulted in almost complete suppression of the enzymatic reaction. Although, this is very likely due to the high azide concentration used, as it would be sufficient to use 1.1 equivalents to still have a concentration significantly higher than K_M. Chemical background observed in the epoxidation of **22** is generally lower compared to **21**, however 200 and 300 mM concentration also showed considerable background reaction lowering the regioselectivity of HheE5 to 70% (Figure 38, B). In all cases, the high azide concentration may have contributed to the overall decrease in regioselectivity and low conversion at

high substrate concentrations. On the other hand, the styrene monooxygenase showed a rather low productivity, hence no further experiments with HHDHs were done to improve regioselectivity at higher than 50 mM epoxide concentration.

Finally, preparative scale reactions using 50 mM (approx. 200 mg) of **21** or **22** coupled with HheE and HheE5 respectively were set up. The catalysts were applied as whole cells and the reactions were carried out in 50 mL reaction volume. HheE could produce 142 mg of **31a** (50% isolated yield) and using HheE5 151 mg of pure **32a** (48 % isolated yield) could be isolated.

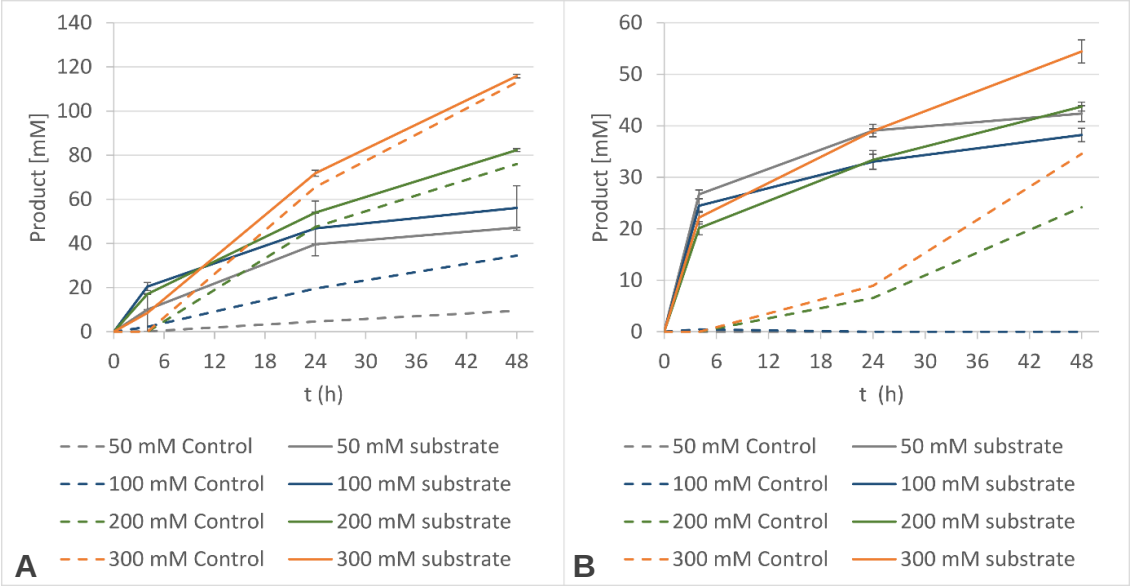


Figure 38. Whole-cell bioconversions of epoxide **21** (graph A) and epoxide **22** (graph B) at different substrate concentrations using HheE and HheE5 (60 g/L, wet cell weight), respectively. Each reaction was carried out in duplicate using the stated substrate concentrations and 2 eq. of NaN_3. Dashed lines represent negative control reactions using *E. coli* cells carrying an empty vector.

3.1.5.3 Compatibility tests

One of the main challenges in setting up cascade reactions is the compatibility of the different reactions. Specific reagents used in a given step can influence the performance of the other catalysts and *vice versa*. Thus, it is important to determine if reactions can proceed in one pot, either at the same time or one after the other, by testing if single compounds from different steps negatively influence another reaction step.

In small scale, no compatibility issues were found for both enzymatic steps as sodium azide present during the epoxidation did not alter StyAB epoxidation (data not shown) hence, epoxidation and epoxide ring opening can run simultaneously or sequentially. Halohydrin dehalogenases could achieve full conversion when added after the

epoxidation step (section 3.1.4) to further validate this conclusion. However, scale-up of the StyAB catalysed epoxidation reaction highlighted the need for a second organic phase to reach reasonably high epoxide concentrations (section 3.1.5.1). Hence, epoxide ring opening catalysed by the HHDHs was tested in the presence of heptane as second phase (25% v/v). Corresponding conversions of racemic *trans*-2,3-epoxyhexane (**21**) and *trans*-2,3-epoxyheptane (**22**) carried out by HheE and HheE5, respectively, are shown in Figure 39. For both enzymes, the presence of the organic solvent had a considerably negative effect on enzyme activity. Especially for HheE, only 33% conversion could be achieved within 24 h (Figure 39, A), whereas HheE5 still reached 87% conversion, however, after significantly longer reaction time compared to the reaction without second organic phase (Figure 39, B). This important negative effect would very likely be one of the challenges to overcome in the setup of this cascade reaction.

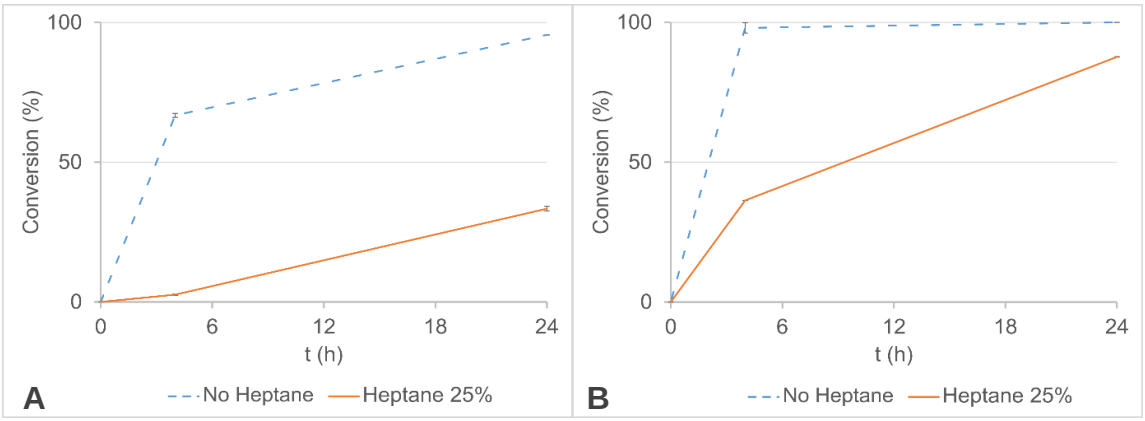


Figure 39. HheE (A) and HheE5 (B) conversion of 10 mM of their respective substrates (**12** for HheE and **22** for HheE5). The blue dashed lines represent control epoxide ring opening reactions without a second organic phase. The orange lines represent conversions in presence of 25% (v/v) heptane. All experiments were done in duplicate.

Similarly, the influence of Pd/C on the epoxidation and epoxide ring opening steps was analysed. In this case, only *trans*-2-heptene (**2**) for the epoxidation reaction and *trans*-2,3-epoxyheptane (**22**) for the epoxide ring opening were used as model substrates. The presence of palladium on carbon caused a decrease in the conversion values for all reactions within 1 h (Figure 40). The biggest effect was observed for the epoxidation reaction when StyAB was employed as whole-cell catalyst (30 g/L wet cell weight) with cellular metabolism for cofactor regeneration. Possible the nanoparticles are toxic to the cells and hence, cofactor regeneration did not work properly. This effect was reduced when BNAH was used for cofactor regeneration and StyA was applied as cell-free extract (4.6 g/L total protein content), confirming again that higher conversions can be achieved in this step when using cell-free extract instead of whole cells. Activity of HheE5, when

employed as whole cell catalyst, was less affected by Pd/C addition, but still a decrease in conversion was evident within 1 h of reaction.

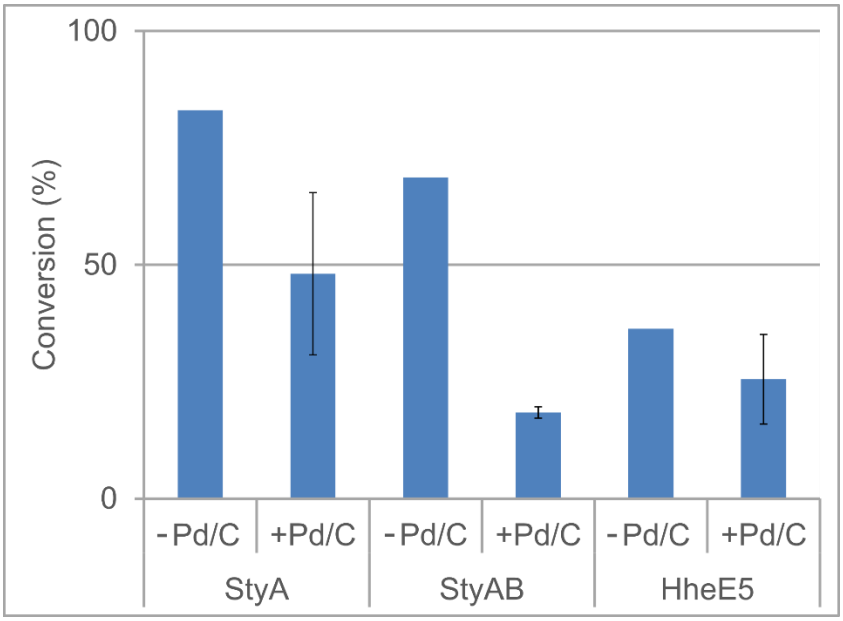


Figure 40. Influence of Pd/C (1 mg/mL) on the different catalysts for step 1 and step 2. All reactions were carried out for 1h using the model substrate were *trans*-2-heptene (**2**) for StyA/StyAB and *trans*-2,3-epoxyheptane (**22**) for HheE5. StyA was used as cell-free extract (4.6 g/L total protein content), with BNAH for cofactor regeneration. StyAB was used as whole cells (30 g/L wet cell weight), with glucose as cofactor regeneration system. HheE5 was used as whole cells (60 g/L wet cell weight).

Finally, the influence of different components from steps 1 and 2 on the azido alcohol hydrogenation was tested. In an optimal cascade setup, Pd/C would be added directly after the epoxide ring opening without workup in between, as the cascade is intended to work in one pot. Hence, a hydrogenation reaction of 5 mM racemic 2-azido-3-hydroxyheptane (**32a**) using 1 mg/mL Pd/C and a hydrogen-containing balloon was performed in aqueous buffer in the presence of single components from step 1 and step 2. Interestingly, all tested components exerted a negative effect on the reaction rate although conversions >90% could still be reached in almost all cases after prolonged reaction time (Figure 41). The most pronounced effect was observed for the cofactors BNAH (10 mM, yellow line) and FAD (200 µM, dark blue line) used by the styrene monooxygenase. Whole cells (60 g/L wet cell weight, green line) had a similar negative effect as the cofactors, whereas the influence of free protein (4.6 g/L, orange line) from the CFE was less and 90% conversion was achieved already after 30 min reaction. The negative influence of by whole cells, BNAH and FAD on Pd/C leaves the option of using both catalysts in step 1 and step 2 as cell-free extract (CFE). In conclusion, ideally the cascade reaction could be run having the first two steps running simultaneously or consequently and adding the hydrogenation catalyst separately after the first two steps

are done. This would avoid having O_2 and H_2 input at the same time and prevent contact between Pd/C and cell-free extract before any azido alcohol product is present.

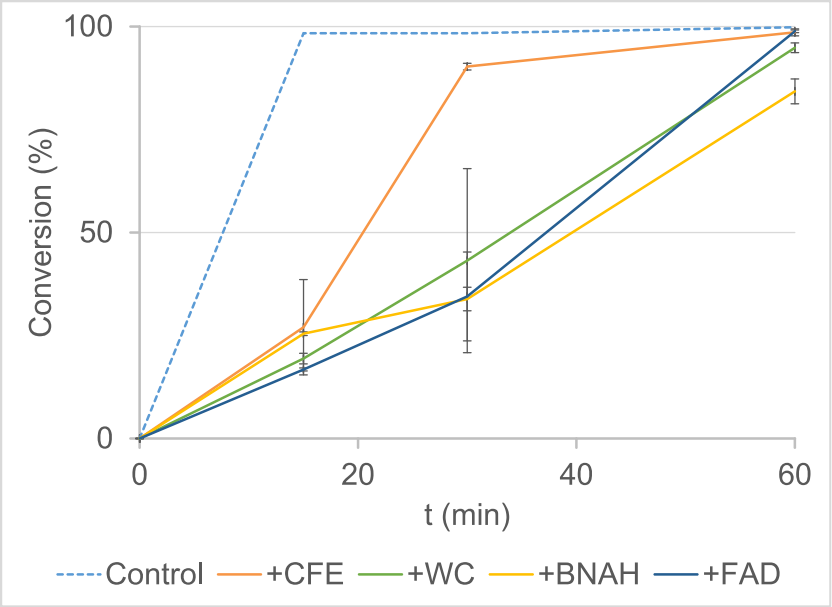


Figure 41. Hydrogenation of 5 mM racemic 2-azido-3-hydroxyheptane (**32a**) using 1 mg/mL Pd/C in aqueous buffer. H_2 was supplied via a balloon connected with a needle. The hydrogenation reaction performed only with Pd/C and H_2 is given as dashed line (control). Components from other cascade steps were added individually: CFE (Cell-free extract 4.6 g/L total protein content, orange line). WC (whole cells 60 g/L wet cell weight, green line). BNAH (10 mM, yellow line) and FAD (200 µM, dark blue line) All reactions were performed in duplicate.

3.1.5.4 Combination of the three reaction steps

After optimization of individual step and cross-interaction studies, the catalysts were combined to perform the cascade reaction for the conversion of 20 mM *trans*-2-heptene (**2**) to (2*R*,3*S*)-2-amino-3-hydroxyheptane (**42a**) as the final product. In detail, the styrene monooxygenase and coexpressed LsADH were applied as cell-free extract using a second organic solvent (heptane 25% v/v) to maximise epoxide production. Giving that the epoxidation reaction is much slower compared to the epoxide ring opening, the first step was run separately for the first 24 h. This delay would allow the production of a critical amount of epoxide and prevent long incubation of the HHDHs with heptane, proven to be toxic for the HHDH. From 24 h to 48 h reaction time, epoxidation and epoxide ring opening occurred in parallel. Using hydrogen for the last step was not possible as the available equipment did not allow a safe and efficient hydrogen handling. For this, alternatives for *in situ* hydrogen generation were explored. Among these, the generation of molecular hydrogen by dropping hydrochloric acid on zinc dust showed an insufficient generation rate under safe conditions. Hydrazine was used for *in situ* H_2 generation in the preparation of authentic compounds (Methods, Section 2.1.1)[158] but did not show any conversion in aqueous solutions and was abandoned. Finally,

triethylsilane (TES) was the only *in situ* hydrogen generation system that gave a reasonable conversion in aqueous buffer. Hence, the final step was performed by adding after the first two steps Pd/C (2 mg/mL) and 3 equivalents of triethylsilane (TES) dropwise for the first 8 h.

Figure 42 presents the progression of the cascade reaction steps over time for the conversion of *trans*-2-heptene (**2**) as model substrate. The hydrogenation reaction was monitored by following the consumption of the corresponding azido alcohol (**32a**), as it was not possible to detect **42a** in GC. According to Figure 42, running the three-step cascade in one pot is feasible although the amino alcohol isolated yield (2.4%) was very low, however compatible with the low azido alcohol produced in the first two steps.

All catalysts reached rather low productivity which added up to the final low yield. In this case, the last two steps appear to be the limiting factors as HheE5 did not fully convert the epoxide produce as well as the hydrogenation catalyst. The epoxide ring opening is negatively influenced by the presence of heptane which cannot be removed as it would significantly decrease the epoxide production.

The hydrogenation could be improved by using an autoclave reactor to provide H_2 overpressure to increase the efficiency of the hydrogenation reaction and fully convert the azido alcohol at least.

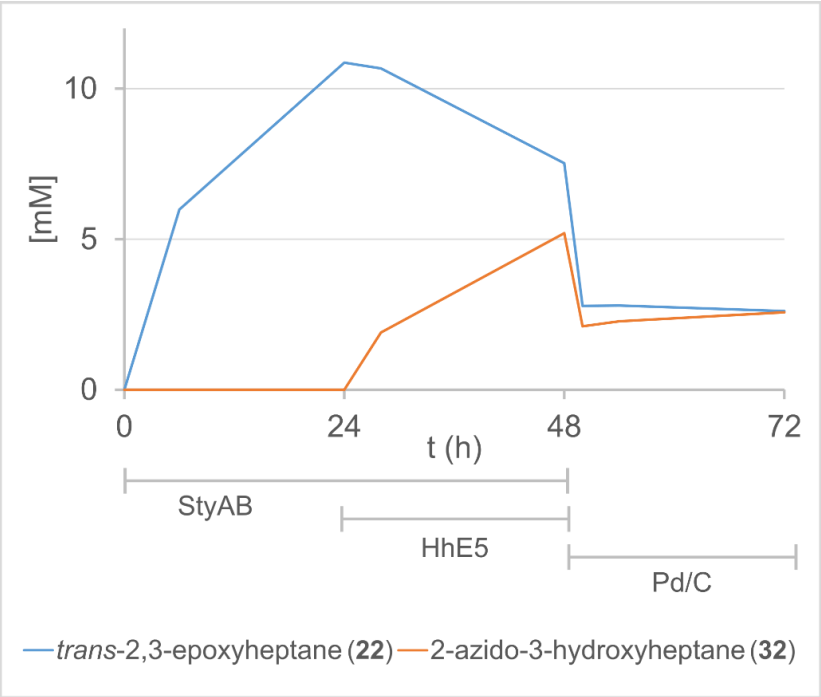


Figure 42. Concentrations (mM) of **22** and **32** produced during the cascade composed of StyAB, HheE5 and Pd/C in 80 mL scale (60 mL + 20 mL heptane as second phase) in the conversion of 20 mM of *trans*-2-heptene. Each catalyst was added sequentially; grey bars show for how long each reaction was performed.

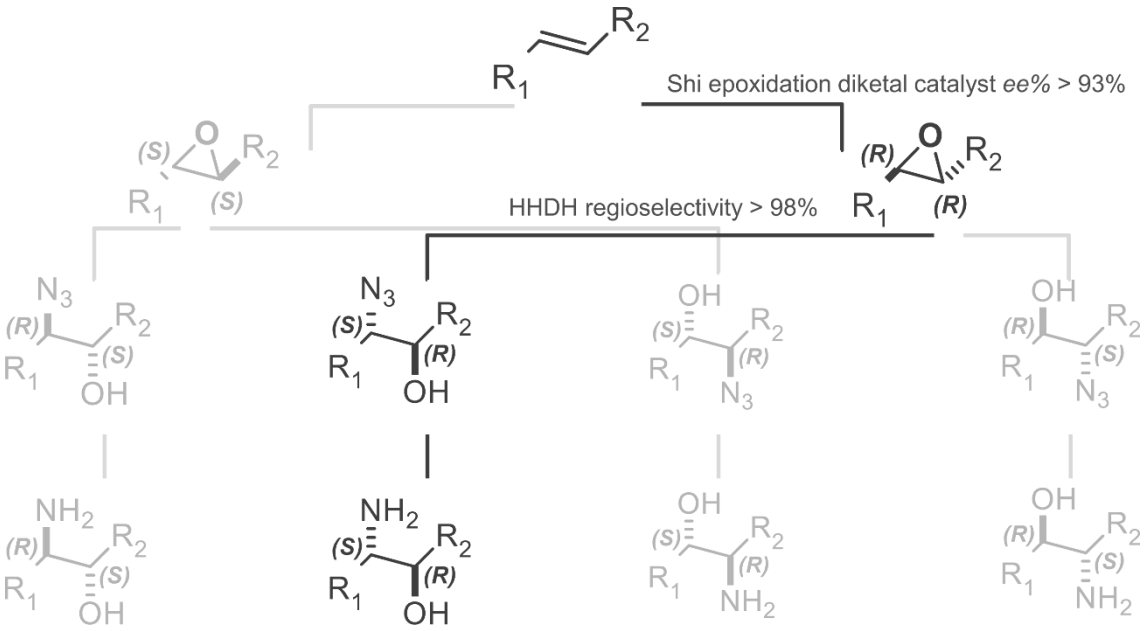
Alternatively, the hydrogenation could be separated by an intermediate work up and run in methanol where full conversion could be reached. Hence, exemplary preparative scale reactions were performed with an intermediate workup before the last (hydrogenation) step. The crude extract was directly dissolved in methanol where hydrazine can be used for an efficient hydrogen generation as used for the synthesis of reference compounds (Methods, Section 2.1.1). Isolated yields for cascade with only the first two steps were determined for comparison. Overall, the reactions where no hydrogenation was performed afforded similar product concentration, however slightly lower isolated yields were obtained. Running the hydrogenation in methanol separately significantly improve the isolated yield compared to the cascade mentioned previously (Figure 42) where only 2.4% isolated yield was reached. Very likely removing free protein significantly improved Pd/C conversion Moreover, extraction of a polar compound from the aqueous buffer is a challenging process and a rather significant amount of product could not be isolated. Very likely the epoxide ring opening step has a strong negative effect as only a maximum of 6.47 mM of **42a** were produced. In all cascade reactions, similar azido alcohol concentrations were reached, although HheE showed a significantly higher sensibility to heptane compared to HheE5 (Figure 39). In all reactions, the corresponding azido alcohols were produced with exquisite enantiomeric excess (> 99%). However, HheE regioselectivity decreased resulting in a regioisomeric ratio of 92:8 (**31a**:**31b**) (entries 1 and 2, a:b, Table 23) probably due to a higher chemical background reaction. Whereas HheE5 displayed a regioselectivity comparable with previous experiments. The high product enantiomeric excess (azido alcohol products) confirmed the validity of this synthetic strategy, yet further optimization is needed to improve the moderate isolated.

Table 23. Summary of the cascade runs combining StyAB, HHDH-catalysed epoxide ring opening and hydrogenation using Pd/C. Entries 1 and 3 summarise the results of the cascades producing (2*R*,3*S*)-2-amino-3-hydroxyhexane (**41a**) and (2*R*,3*S*)-2-amino-3-hydroxyheptane (**42a**) respectively. Entries 2 (production of **31a**) and 4 (production of **32a**) summarise the results of the cascades without hydrogenation step. In all cascades, 50 mM of the respective alkenes **1** (entries 1 and 2) an **2** (entries 3 and 4) were used as substrate.

#	Cascade	2-azido- 3-alcohol (mM)	Theoretical yield (mg)	Isolated Yield (mg)	Isolated Yield [%]	ee (epoxide)	ee | a:b (azide)
1	StyAB, HheE and Pd/C	5.67	39.89	29.8	8.5	95%	>99% | 92:8
2	StyAB, HheE	6.14	52.76	31.7	7.0	95%	>99% | 92:8
3	StyAB, HheE5 and Pd/C	6.16	48.52	40.9	10.5	98%	>99% | 99:1
4	StyAB, HheE5	6.47	61.08	42.9	9.1	98%	>99% | 99:1

3.1.6 Shi epoxidation diketal catalyst – HHDH – Pd/C cascade

Scheme 20 illustrated that regioselective epoxide ring opening of epoxides **21** and **22** in *S*,*S*-configuration results in the production of the corresponding enantiopure azido- and amino alcohol **31a** and **32a** in (2*R*,3*S*)-configuration. Replacing the StyAB of the first step with a catalyst producing the (*R*,*R*)-epoxide enantiomers but keeping the same HHDH for the second step would result in a complementary cascade for the synthesis of enantiocomplementary azido- and aminoalcohols (Scheme 21). Thus, combining the Shi epoxidation catalyst with HheE or HheE5 and Pd/C would result in the production of (2*S*,3*R*)-2-azido-3-hydroxyhexane (**31a**) and (2*S*,3*R*)-2-azido-3-hydroxyheptane (**32a**) respectively. As for the cascade described in 3.1.5, single steps were optimized separately and compatibility tests were run to determine the best way to combine the catalysts.



$R_1 = CH_3$; $R_2 = (C_2H_2)_2CH_3$, $(C_2H_2)_3CH_3$

Scheme 21. Example of the cascade composed of the (*R*)-selective Shi epoxidation diketal catalyst in the first step, a regioselective halohydrin dehalogenase in the second step and palladium on carbon in the last step. As seen for cascade 1 in Scheme 20, the combination of enantioselectivity and regioselectivity leads to the formation of only one of the four possible product stereoisomers.

3.1.6.1 Optimization of Shi epoxidation

The Shi epoxidation diketal catalyst is an organocatalyst and in particular, it is a chiral ketone. According to the reaction mechanism, the keto group reacts with Oxone© to form a dioxirane moiety, necessary to perform the epoxidation. The surrounding of the keto

group must be chiral to carry out an asymmetric epoxidation.[187] In the first small scale reaction, 10 mM *trans*-2-heptene (**2**) concentration was used and the epoxidation reaction was run for 24 h keeping the pH at 8 by adding $NaHCO_3$. In this case, the product enantiomeric excess obtained was only 60% (data not shown). Shi *et al.* determined that in some cases performing the reactions at low temperature (-2 °C) and at pH>10 could help to prevent the catalyst decomposition.[187] With these reaction conditions it was possible to yield (*R,R*)-*trans*-2,3-epoxyheptane (**22**) in 94% product enantiomeric excess (Figure 15, Section 3.1.1.4). In all small scale reactions, the acetonitrile concentration used was 20% (v/v), however the amount needed may vary depending on the substrate reactivity, hence different concentrations were tested. Additionally, both HHDHs showed sensibility to the presence of a second organic phase (section 3.1.5.3), hence lower acetonitrile concentrations were tested (5 and 10% v/v). The reaction was first optimized in 50 mL scale as Oxone© should be added portion wise. Giving that Oxone© is a powder, a larger scale reaction allows the experimenter to use a spatula to add it in a reproducible way. 50 mM substrate concentration and 0.2 equivalents of the catalyst were used. Up to four equivalents of Oxone© were added to the reaction in portions. For both substrates **1** and **2** a 20% (v/v) concentration of acetonitrile resulted in the best conversion values. The product enantiomeric excess decreased to 80-85% compared to the high ee_P obtained in small scale reactions. Overall, better conversions were obtained in the epoxidation of t*rans*-2-hexene (**1**) compared to when *trans*-2-heptene (**2**) was used as substrate (Figure 43). Decreasing the amount of acetonitrile resulted in significantly lower conversions using both substrates. Lowering acetonitrile concentration to 10% (v/v) caused a decrease in conversion below 50% for both substrates, however the product enantiomeric excess remained around to 80-85% (Figure 43). Interestingly a further decrease of acetonitrile to 5% (v/v) did not cause any decrease in conversion, although the ee_P decreased to approx. 50%. Very likely, Oxone© could better dissolve in solution at lower acetonitrile concentrations causing auto-oxidation of the Shi diketal epoxidation catalyst leading to the formation of acetone. Acetone can perform the same epoxidation cycle, however, displaying no enantioselectivity. As mentioned previously, acetonitrile is proposed to participate in the reaction mechanism, hence lowering its concentration resulted in lower yields in agreement with this hypothesis.

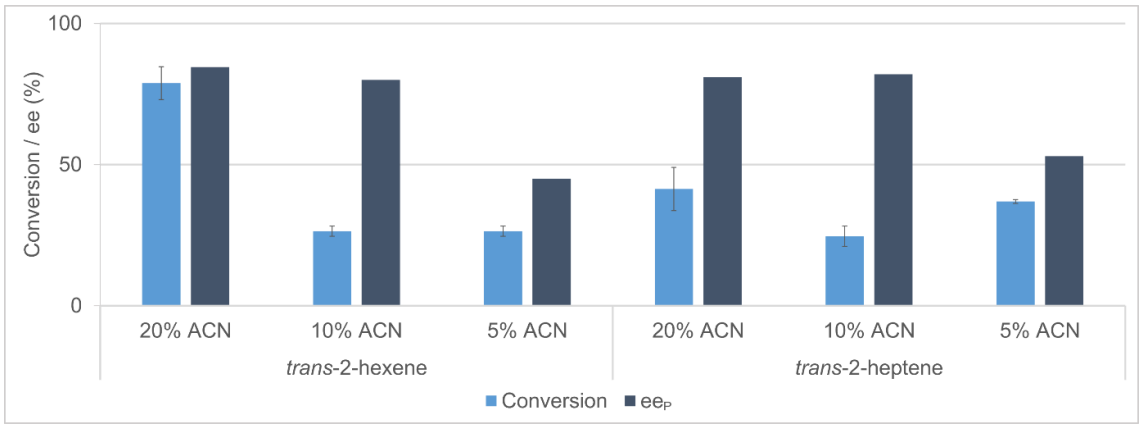


Figure 43. Conversion and product enantiomeric excess (ee_P) obtained at different acetonitrile (ACN) concentrations. The epoxidations were carried out using 0.2 equivalents of the Shi epoxidation catalyst and four equivalents of Oxone in the conversion of 50 mM of either **1** or **2**.

A higher substrate concentration (250 mM) was tested to investigate if the reaction could be upscaled without additional challenges. Furthermore, this way the ratio Oxone©:acetonitrile would increase similarly to reactions with only 5% acetonitrile concentration. Hence, a decrease of ee_P would reveal if an excess of Oxone© compared to acetonitrile is causing the lower ee_P or acetonitrile alone. Figure 44 shows that although it was possible to produce 6 g/L (59 mM) of *trans*-2,3-epoxyhexane and 8 g/L (69 mM) of *trans*-2,3-epoxyheptane the overall conversion was below 50%. The product enantiomeric excess did not decrease as observed for lower acetonitrile concentrations (Figure 43). Thus, Oxone© seems to have only a minor effect on the enantioselectivity of the reaction. Overall, the product concentrations reached in the Shi epoxidation are higher compared to reactions with StyAB, although the products are formed with a lower product enantiomeric excess.

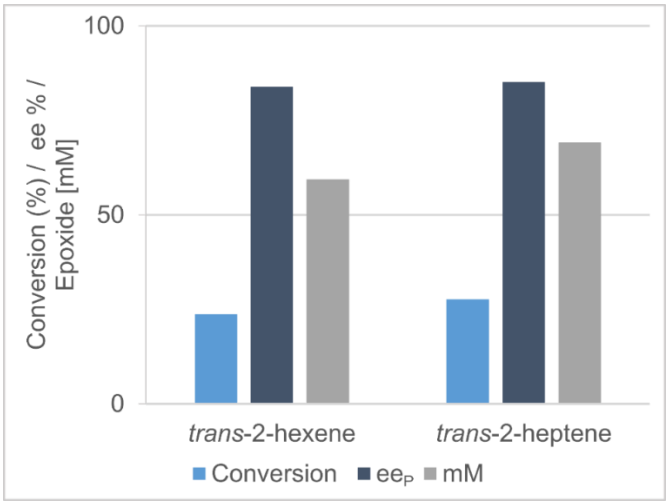


Figure 44. Conversion (light blue), product enantiomeric excess (ee_P, dark blue) and final epoxide concentration (grey) obtained for the epoxidation of 250 mM of alkene substrates (**1**) and (**2**). In all cases, 0.2 eq. of Shi diketal catalyst, acetonitrile to a concentration of 20 % (v/v) and 4 eq. of Oxone added portion wise were used. Both reactions were carried out at -2 °C.

3.1.6.2 Compatibility tests

As described in section 2.2.4, the Shi epoxidation diketal catalyst requires extreme reaction conditions; in fact, a temperature below zero is required together with a pH of 11 to ensure catalyst stability. Hence, this cascade must be run in sequential mode, as the high pH and low temperature are not compatible with the optimal reaction conditions required by halohydrin dehalogenases. Furthermore, the epoxidation reaction requires a significant amount of acetonitrile (Figure 43) to reach good conversions and to preserve the enantioselectivity of the catalyst. Considering that HheE and HheE5 activity was significantly affected by the solvent heptane, the effect of acetonitrile on both enzymes was studied as well. Similarly, the influence of other components present in the Shi epoxidation was investigated on the epoxide ring opening of **12** or **22** using HheE and HheE5, respectively (Figure 45). Again, both HHDHs exhibited a sensibility to acetonitrile as the activity was already reduced substantially at 10% acetonitrile concentration within the first 4 h. 5% acetonitrile had a significantly lower effect on HheE and no differences with the control reaction were observed for HheE5. In the presence of 20% acetonitrile, however, HheE5 activity was almost abolished. These reactions were carried out using whole-cell biocatalysts (60 g/L wet cell weight), but additionally, a reaction using HheE5-containing cell-free extract was performed in the presence of 5% acetonitrile. Whole cells yielded a slightly better conversion indicating that the cellular envelope probably protects the enzyme against the organic solvent improving the conversion.

Additionally, two different concentrations of Oxone© (50 mM and 10 mM) were tested in combination with HheE5 to check if any unreacted compound from step 1 (epoxidation) could be an issue for the HHDHs. The respective graphs in Figure 45 indicate that higher concentrations of Oxone© significantly affected HheE5 activity and only 27% conversion could be achieved after 24 h in the presence of 50 mM Oxone©. However, significant foaming was observed after NaN_3 addition, Oxone© can oxidise the azide to form N_2 and N_2O in ratio 2:1,[188] resulting in a lower concentration of the nucleophile for epoxide ring opening. When only 10 mM Oxone© was added, substrate conversion was reduced less. In this case, if all Oxone© would have reacted with azide in a 1:1 ratio, a theoretical amount of 30 mM azide (3 eq.) would still have been available. Moreover, it appeared that Oxone© had a greater effect on HheE5 than on HheE, for which the presence of Oxone© even seems to be beneficial as a slightly higher conversion was reached after 4 h compared to the positive control reaction without Oxone© addition.

In summary, the presence of acetonitrile will be the limiting factor for cascade setup as seen for cascade 1 (section 3.1.5.3), especially in the case of HheE. Moreover, the effect of Oxone© on HheE5 at high concentration would be less critical if step 2 (epoxide ring opening) would be started after completion of the first step (epoxidation), once all added

Oxone© has either reacted with the Shi epoxidation diketal catalyst or spontaneously decomposed.[189]

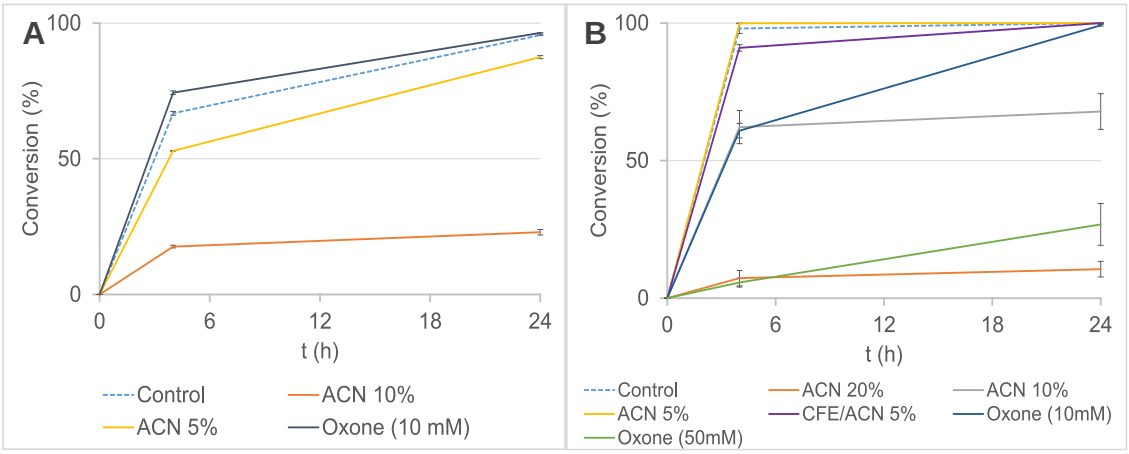


Figure 45. Epoxide ring opening catalysed by HheE (A, 10 mM of **12** as substrate) and HheE5 (B, 10 mM of **22** as substrate) using whole cells (60 g/L wet cell weight) or cell-free extract (purple line in graph B, 9.2 g/L total protein content) and NaN_3 (40 mM) in presence of different components of the Shi epoxidation reaction. All reactions were performed in duplicate. Control reactions without acetonitrile or Oxone© are given as dashed lines.

As mentioned previously, cascade reactions often require a compromise between optimal conditions of the separate reaction steps. According to Figure 43, the best acetonitrile concentration for the Shi epoxidation is 20%. In contrast, due to the sensibility of the selected HHDHs to organic solvents, maximum 5% acetonitrile should be used.
All in all, this cascade should be run as a 1-pot-3 steps cascade reaction giving the extremely different reaction conditions of the first two steps. Moreover, as shown in Figure 41, the hydrogenation catalyst is negatively affected by the presence of whole cells. Hence, an intermediate workup to remove whole cells from the reaction mix before the hydrogenation may be necessary resulting in a 2-pot reaction.

3.1.6.3 Combination of the three reaction steps

Finally, as for cascade 1, a preparative cascade run on 50 mL scale using an acetonitrile concentration of 5% was performed for the conversion of 50 mM *trans*-2-heptene (**2**) to (2*S*,3*R*)-2-amino-3-hydroxyheptane (**42a**) (Figure 46). The epoxidation step was run first for 24 h, as the required low temperature (-2 °C) and the high pH (11) did not allow a direct combination with HheE5. Hence, after heating the reaction mix to 25 °C and adjusting the pH to 7.5, HheE5 (60 g/L, wet cell weight) and sodium azide (50 mM) were added and the epoxide ring opening was carried out for additional 24 h. At last, Pd/C (2 mg/mL) and 3 equivalents of TES were added stepwise for 8 h, as mentioned previously. The progress of the three reaction steps is displayed in Figure 46. The reaction was

followed by formation of epoxide **22** and azido alcohol **32**, whereas the hydrogenation step was followed by consumption of **32**. In general, this 1-pot-3-step cascade appeared feasible, however, the first step seemed to be the bottleneck yielding only a conversion of 33% and a low ee_P of 34% in accordance with previous experiments (Figure 43). HheE5 performed better in this cascade with only 5% acetonitrile concentration, whereas in other cascade runs using 10% or 20% acetonitrile concentration no further conversion could be observed in the epoxide ring-opening step (data not shown). Compared to cascade 1, a higher conversion was achieved in the hydrogenation step reaching 64% after 24h and the final isolated yield (3%) was slightly better than cascade 1 (2.4%). In conclusion, despite the rather low ee_P obtained, this process with further process engineering to overcome the acetonitrile limitation could result in the production of enantiopure molecules.

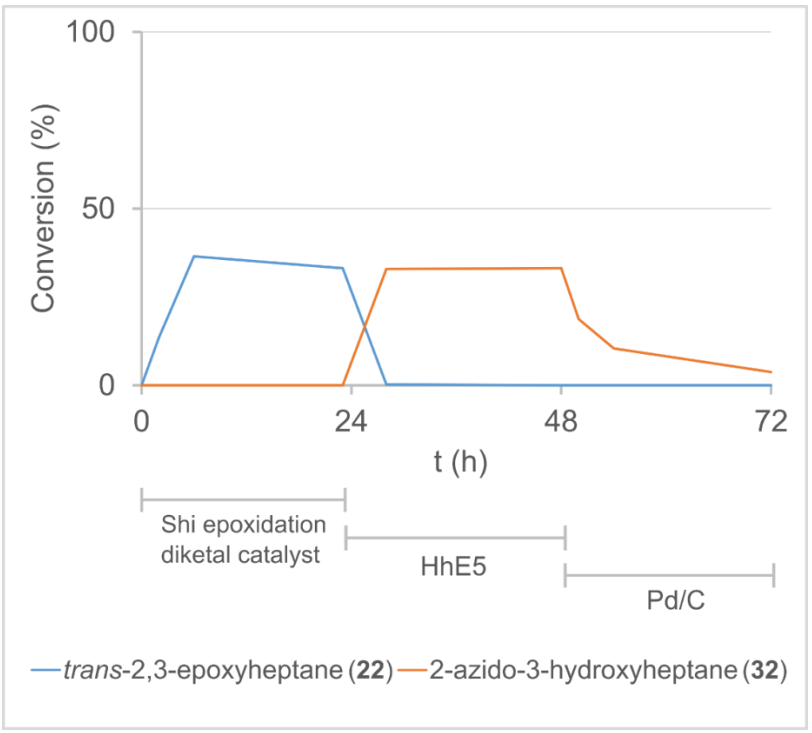


Figure 46. Conversion values of the sequential cascade composed of Shi epoxidation, HheE5-catalysed epoxide ring opening and hydrogenation using Pd/C in 50 mL scale (5% acetonitrile concentration) starting from 50 mM *trans*-2-heptene. The epoxidation and epoxide ring opening reactions were monitored via GC, the hydrogenation was monitored via **32** consumption.

Similarly to cascade 1, preparative scale reactions were set up separating the hydrogenation reaction from the first two steps (2-pot-3-step cascade). This generally improved the efficiency of reaction work up as amino alcohols are polar compounds and are typically difficult to extract from the aqueous phase. Hence, performing the

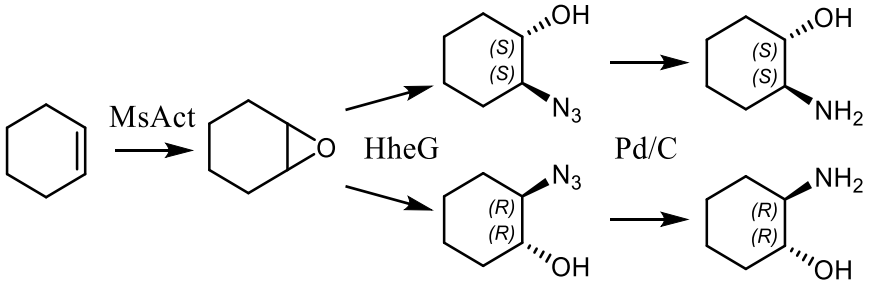
hydrogenation step in methanol could improve the overall yield as no amino alcohol extraction from the aqueous phase was needed. Entries 1 and 3 in Table 24 display the isolated yields of the cascade for the production of (2*S,*3*R*)-2-amino-3-hydroxyhexane (**41a**) and (2*S,*3*R*)-2-amino-3-hydroxyheptane (**42a**), respectively, whereas entries 2 and 4 summarise the data corresponding to the 2-step cascades including only epoxidation and epoxide ring opening. Despite the lower kinetics, the cascades involving HheE achieved similar azide concentrations. For all cases, the products were obtained with rather low ee_P as a direct result of the poor enantioselectivity achieved in the epoxidation step. Enantiomeric excesses were further improved in the HHDH-catalysed step as both enzymes preferentially converted the epoxide in (*R,R*)-configuration. Moreover, lower regioselectivity is observed for entry 1, which might be caused by a higher chemical background reaction. However, as discussed before, the two regioisomers can be separated by silica column chromatography (Appendix, Supporting figure 25). Compared to cascade 1 higher isolated yields were achieved, however with significantly lower ee_P. Increasing the acetonitrile concentration in the first step would further increase conversion and highly increase the ee_P of the overall cascade. However, further process engineering would be required to efficiently remove or decrease the presence of acetonitrile before starting the epoxide ring opening reaction.

Table 24. Summary of the cascade runs combining Shi epoxidation, HHDH-catalysed epoxide ring opening and hydrogenation using Pd/C. Entries 1 and 3 summarise the results of the cascades producing (2*S,*3*R*)-2-amino-3-hydroxyhexane (**41a**) and (2*S,*3*R*)-2-amino-3-hydroxyheptane (**42a**), respectively. Entries 2 (production of **31a**) and 4 (production of **32a**) summarise the results of the cascades without hydrogenation step. In all cascades, 50 mM of the respective alkenes **1** (entries 1 and 2) and **2** (entries 3 and 4) were used as substrate.

#	Cascade	2-azido- 3-alcohol (mM)	Theoretical (mg)	Isolated Yield (mg)	Isolated Yield	ee% (epoxide)	ee% | a:b (azide)
1	Shi, HheE and Pd/C	24	157.5	86.1	26%	42%	60% | 88:12
2	Shi, HheE	19.8	158.77	60.5	15%	45%	61% | 97:3
3	Shi, HheE5 and Pd/C	15.45	113.53	67.5	18%	53%	58% | 98:2
4	Shi, HheE5	12.35	108.73	83.9	19%	47%	56% | 99:1

3.2 Cascade starting from cyclic alkenes

The main difference of this cascade compared to the cascades described in section 3.1 is the use of a cyclic alkene, cyclohexane, as starting substrate. Epoxidation of this compound leads to a prochiral epoxide, hence, the strategy in this cascade was slightly different. Enantioselectivity is not required in the first step but in the subsequent epoxide ring opening. Thus, MsAct was selected for its high perhydrolysis/hydrolysis ratio in water to induce the chemical epoxidation (step 1), whereas HheG was selected as the catalyst to perform the stereoselective epoxide ring opening (step 2), as this HHDH was the only known enzyme with significant activity towards this substrate.[105] Finally, palladium on carbon was used for the last step to reduce the azido group and to produce the final enantiopure amino alcohol (Scheme 22).



Scheme 22. Reaction scheme for cascade 3 starting from cyclohexene combining epoxidation in the first step, azidolysis of the corresponding epoxide in the second step and chemical hydrogenation of the azido group in the last step to form an amino alcohol. Stereoselectivity is required in the second step to generate an enantiopure product.

3.2.1 Epoxidation

3.2.1.1 Expression and purification of the acyl transferase from *Mycobacterium smegmatis* (MsAct)

The heterologous expression of the acyl transferase was established previously.[161] pET28a(+) was used as expression vector and *E. coli* BL21 (DE3) was chosen as the host organism for heterologous production of the N-terminally His-tagged protein. MsAct was purified using the IMAC purification protocol. However, the standard imidazole gradient resulted in a slow elution of the target protein as shown by the broad MsAct peak in the chromatogram in Figure 47. To limit the further dilution of the sample, imidazole concentration was set to 500 mM around fraction 47 (Figure 47). This concentration increase resulted in a prompt elution of MsAct, which is reflected by a sudden increase in the absorbance at 280 nm as well as a stronger protein band in the corresponding SDS-PAGE gel starting from fraction 48. After concentration of the

combined elution fractions by ultrafiltration using a 30 kDa cut-off membrane, a yield of 160 mg/L of pure MsAct was obtained.

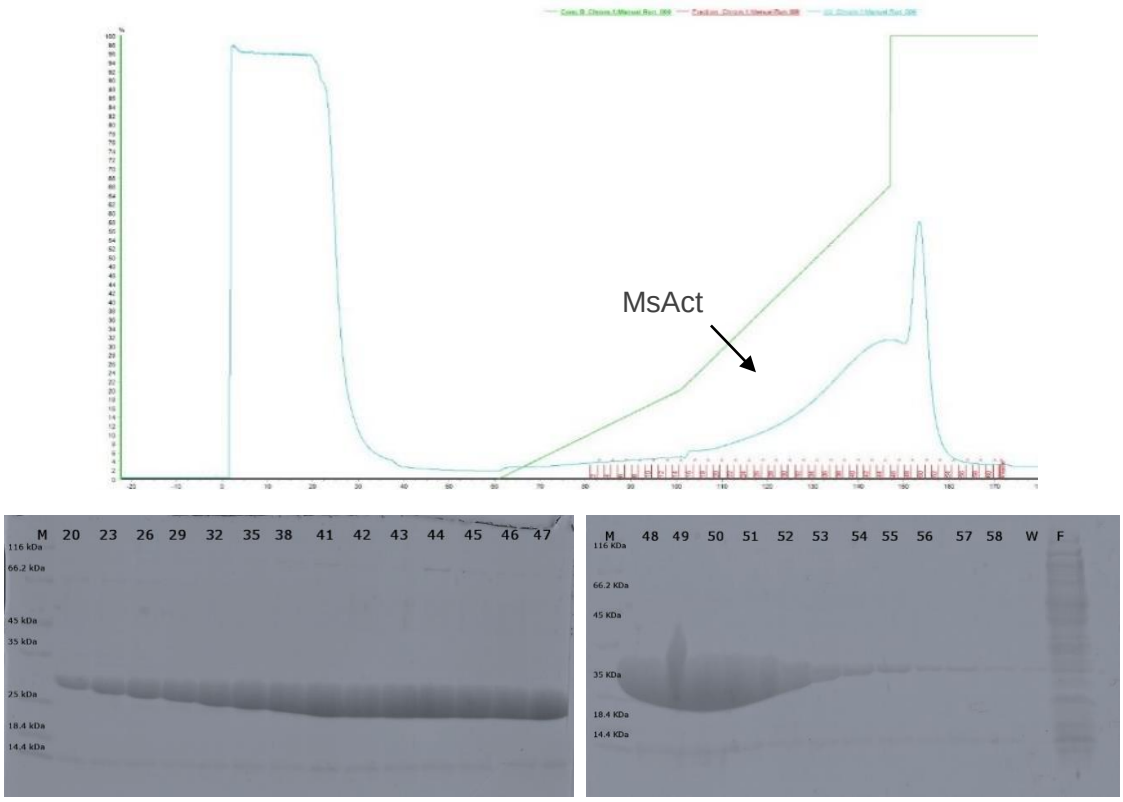
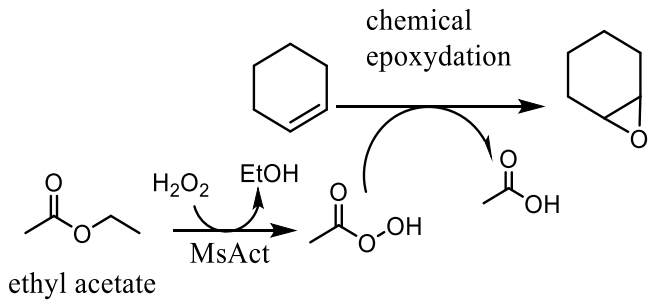


Figure 47. The upper picture shows the purification chromatogram of the His-tagged MsAct. UV signal representing the protein elution is given in blue, whereas the gradient progression is represented by the green line. In the SDS-PAGE gels, only the fractions containing the target protein are shown.

3.2.1.2 MsAct-catalysed perhydrolysis and subsequent chemical epoxidation



Scheme 23. General reaction scheme of the epoxidation of cyclohexene catalysed by MsAct. The enzyme catalyses the perhydrolysis of ethyl acetate using hydrogen peroxide. The resulting peracid spontaneously reacts with cyclohexene producing cyclohexene oxide.

As shown in Scheme 23, the actual role of the acyl transferase from *Mycobacterium smegmatis* in the first reaction step of this cascade is the generation of a peracid (peracetic acid in this specific case). The peracid promotes chemical epoxidation of cyclohexene resulting in the formation of cyclohexene oxide, an achiral epoxide. Hence, stereoselectivity is not required in this step. Several lipases have also been shown to catalyse the perhydrolysis of esters, however, usually their hydrolysis activity is significantly higher in the presence of water. Hence, perhydrolysis using lipases is often performed in an organic solvent such as ethyl acetate to reduce hydrolysis.[190] In contrast, MsAct has an extremely hydrophobic cavity resulting in a preference of the enzyme for perhydrolysis over hydrolysis. This way, the reaction using MsAct can be run in aqueous buffer in the presence of ethyl acetate as substrate. This characteristic made MsAct a promising candidate for the combination of this epoxidation reaction with an epoxide ring opening catalysed by HheG. Next to the ester, the epoxidation reaction requires also the presence of hydrogen peroxide for the enzyme to produce the peracid. Three different concentrations of ethyl acetate (10, 20 and 30% v/v) and different amounts of H_2O_2 (1, 2, 3, 4, 10, 4x2.5 equivalents) were tested for the efficient epoxidation of 10 mM cyclohexene. As a result, 10% ethyl acetate resulted in significantly better conversions regardless of the amount of hydrogen peroxide used (Figure 48). 20% and 30% ethyl acetate showed very low conversion hinting at a possible enzyme inactivation caused by the solvent or the acid produced. Using a larger excess of hydrogen peroxide yielded higher product concentration. Furthermore, it appeared that adding hydrogen peroxide in portions was a better strategy (Figure 48). Possibly, hydrogen peroxide degrades over time and MsAct cannot effectively use all the supplemented H_2O_2 from the beginning, hence adding H_2O_2 in portions increased the total amount of peroxide that can be used by the enzyme.

Additionally, acetic acid is produced during the reaction as by-product of the chemical epoxidation (Scheme 23) and by hydrolysis of ethyl acetate, although the latter is less probable giving the enzyme's preference to catalyse perhydrolysis.

Changing the buffer to saturated $NaHCO_3$ solution to keep the pH at 8 resulted in high substrate and product evaporation due to the formation of CO_2 and no product could be detected in GC-FID (data not shown). Very likely, a tight pH control could improve the overall reaction by balancing the excess of acetic acid formed during the chemical epoxidation and by the enzyme. Additionally, an optimal feed of hydrogen peroxide could further improve the obtained yields. This, however, was not further explored, as no HheG variant with sufficient enantioselectivity in the ring opening of cyclohexene oxide could be obtained during HheG engineering. Hence, also the combination of the two reaction steps was not tested.

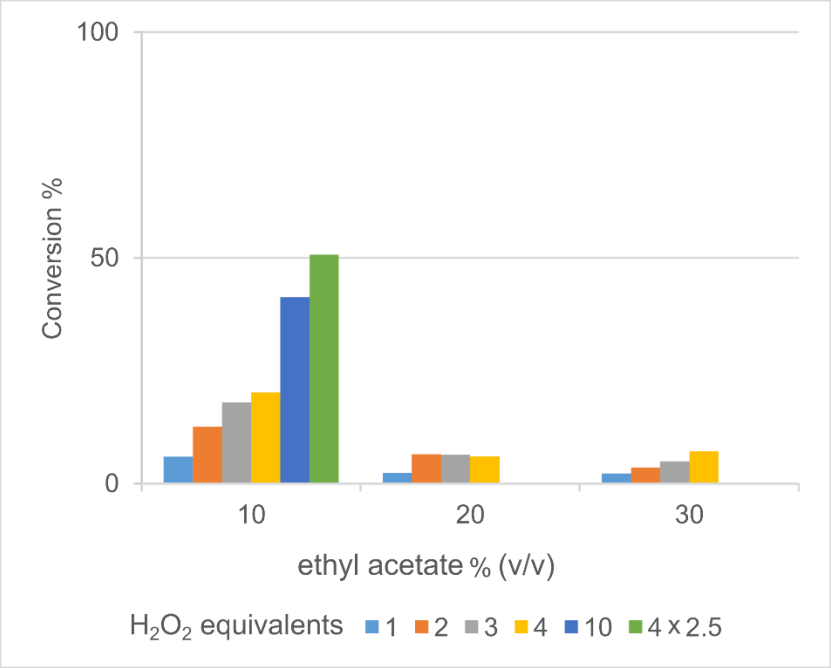


Figure 48. Epoxidation of cyclohexene (5 mM) promoted by MsAct in the presence of ethyl acetate at different concentrations (% v/v) and using different amounts of H_2O_2 (1, 2, 3, 4, 10, 4x2.5 equivalents).

3.2.2 Epoxide ring opening

3.2.2.1 In *silico* mutagenesis of HheG using the Rosetta software suite

HheG was previously found to be the only HHDH with significant activity towards cyclohexene oxide.[105] However, it formed the corresponding azido-cyclohexanol with only 40% enantiomeric excess. Hence, an in *silico* design was attempted to improve HheG´s enantioselectivity using the "coupled moves" application in the Rosetta framework.[149] This application efficiently introduces backbone as well as ligand movements in the design resulting in a significantly higher accuracy compared to the previous "fixed backbone" method. The coupled moves application requires a substrate bound PDB file and a list with the residues to be included in the *in silico* mutagenesis. Since no substrate-bound structure of HheG was available, cyclohexene oxide or (1*R*,2*S*)-2-cyanocyclohexanol were manually docked in HheG's crystal structure (PDB-ID: 5O30) after aligning it with the corresponding substrate-bound structures of HheC (PDB-ID: 1ZMT and 4IXT). For the design, all HheG residues within 5 Å from the ligand excluding the catalytic triad were included in the virtual mutagenesis (14 positions). The typical output of the application is a set of 20 FASTA files (one for each of the 20 runs performed) including all sequences that carry one or more mutations improving the protein-ligand interaction. The sequences were pooled and filtered for redundancy ending with approximatively 2000 unique amino acid sequences. The variability of the resulting sequences can then be visualized using sequence logos.

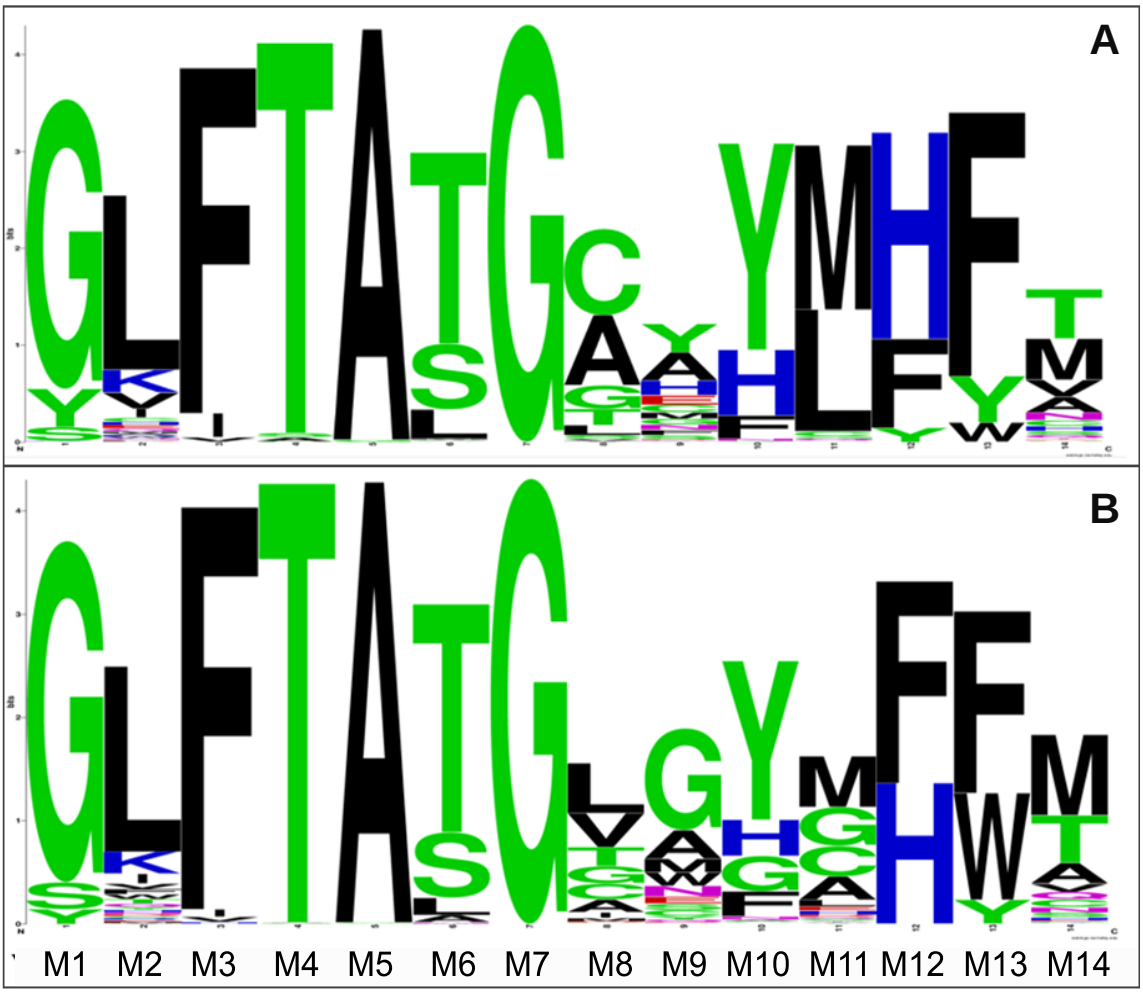


Figure 49. Visualization using sequence logos of the two *in silico* mutagenesis with either cyclohexene oxide (A) or (1*R*, 2*S*)-2 cyanocyclohexanol (B) in the active site of HheG crystal structure (5O30).

In Figure 49, sequence logos of two distinct runs are shown starting from HheG structures with cyclohexene oxide (A) and 2-cyanocyclohexanol (B) docked as ligands. The corresponding 2-azidocyclohexanol could not be used as ligand due to a software limitation as the azido group is not yet parametrised. Excluding the catalytic triad, 14 residues within 5 Å distance from the ligand were randomized in HheG during the Rosetta run. Among these, six positions showed high variability in the resulting sequence set, as visible for positions M2, M8 to M11 and M14 in Figure 50. This indicates that the wild type amino acids in these positions were not the best option for optimal ligand binding. Of these, positions M2, M8, M9 and M14 were selected for mutagenesis including only the top three highest scoring amino acids in each position. All the desired mutations were introduced alone or in combination using site-directed mutagenesis. Resulting mutants were tested for their activity and enantioselectivity in epoxide ring opening reactions of cyclohexene oxide using azide as nucleophile. Most of the mutations resulted in slower epoxide ring opening and only a few mutants could reach 100% conversion in the first hour of reaction as found for the wild type (Figure 49, A).

The latter mutants (MUT2, 6, 12, 14 and 22), however, displayed similar or worse enantioselectivity compared to the wild type (Figure 49, B). MUT23 on the other hand, formed the azido alcohol with a significantly enhanced enantiomeric excess (from 45% ee for the wild type to 71% for the mutant), however, with lower conversion rate. Surprisingly, five of the 24 generated mutants displayed a reversed selectivity (MUT8, 9, 15, 16, 17, 18). In particular, MUT18 yielded 2-azidocyclohexanol with 66%ee after one hour but for the opposite enantiomer compared to the wild type (data not shown). This enantiomeric excess decreased to 37% after 22 h reaction time due to the high chemical background reaction and the low enzymatic activity.

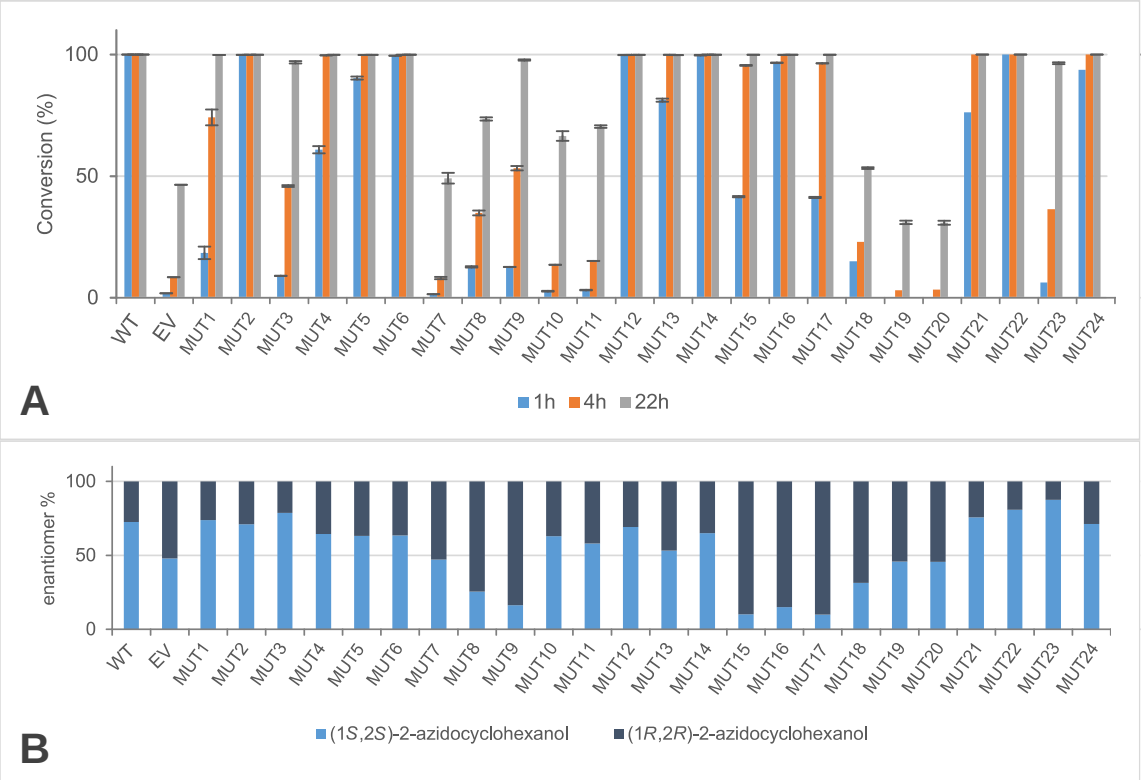


Figure 50. (A) Conversions of cyclohexene oxide with NaN_3 as nucleophile catalysed by the different HheG mutants and (B) ratio of the different enantiomers produced. The reactions were carried out using whole cells, 20 mM cyclohexene oxide and 40 mM NaN_3. EV = negative control using *E. coli* BL21 (DE3) cells containing empty vector. Enantiomer ratios are shown for the 22h reaction sample.

Unexpectedly, using NaCN as nucleophile in the screening of the HheG variants produced quite different results. First of all, some of the mutants displayed such a low activity that it hardly exceeded the chemical background reaction. Hence, in such cases it was necessary to generate cyclohexene oxide *in situ* starting from 2-chlorocyclohexanol as this resulted in lower chemical background in the cyanolysis of cyclohexene oxide. Moreover, most of the mutants were less active than the wild type

(Figure 51), which is not in agreement with the results obtained for the azidolysis reaction (Figure 50). Only variant MUT22 displayed higher conversion than wild-type HheG. Additionally, no improvement in the enantioselectivity was observed for any of the mutants (Figure 51, B), which contradicts the results obtained for the same mutants in the azidolysis of cyclohexene oxide. This seems to indicate that the enantioselectivity of HheG in the ring opening of cyclohexene oxide is dependent on the type of nucleophile used. As an example, MUT15 displayed inverted enantioselectivity in the azidolysis of cyclohexene oxide and a conversion similar to wild type, whereas no activity could be observed when cyanide was used as nucleophile.

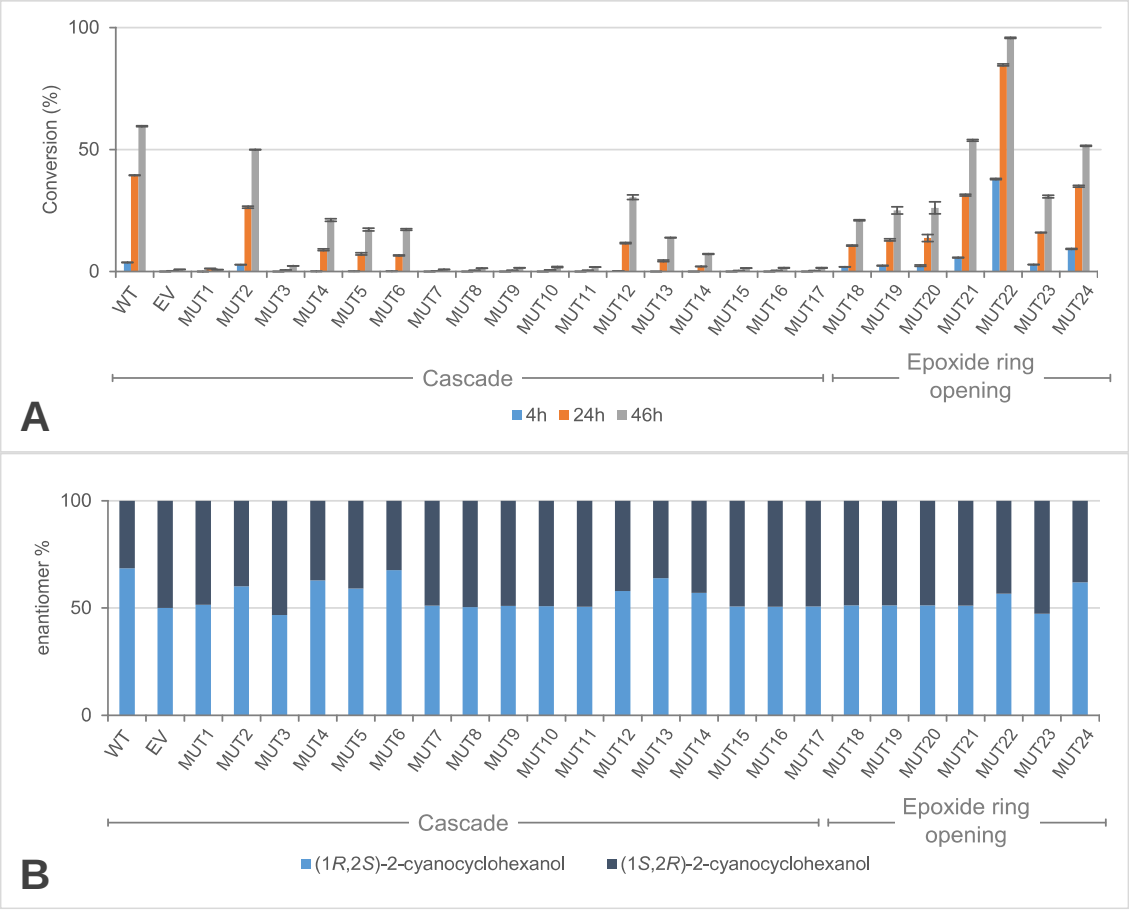


Figure 51. (A) Conversions of cyclohexene oxide with NaCN as nucleophile catalysed by the different HheG mutants and (B) ratio of the different enantiomers produced. The reactions were carried out using whole cells, either 20 mM 2-chlorocyclohexanol (cascade) or 20 mM cyclohexene oxide (epoxide ring opening) and 40 mM NaN_3. EV = negative control using *E. coli* BL21 (DE3) cells containing empty vector. Enantiomer ratios are shown for the 46 h reaction sample.

In conclusion, the *in silico* mutagenesis of HheG could correctly predict positions with a significant effect on enantioselectivity, however, activity of the mutants was lower than for wild type in most cases. Interestingly, a few mutants displayed even inverted

enantioselectivity compared to wild type in the azidolysis of cyclohexene oxide. Unfortunately, no HheG variant producing 2-azidocyclohexanol with >90%ee could be identified. Hence, the combination of MsAct, HheG and Pd/C in one cascade was not tested.

4. Discussion

In this work, a novel chemo-enzymatic cascade for the production of enantiopure vicinal amino alcohols from non-terminal alkenes was established. The synthetic route comprised three steps: 1) epoxidation of the alkene, 2) epoxide ring opening with azide as nucleophiles and 3) hydrogenation of the azido group. Stereoselectivity in the first step and regioselectivity in the second step enabled control of the stereoconfiguration of the two chiral centres in the final amino alcohol. The styrene monooxygenase from *Rhodococcus sp.* ST-10 and the Shi epoxidation diketal catalyst were selected to perform the epoxidation as they selectively synthesise opposite epoxide enantiomers. HheE and HheE5 were selected for the regioselective azidolysis of epoxides **21** and **22**, respectively. Lastly, Pd/C was used to catalyse the reduction of the azido alcohols to the corresponding amino alcohols.

The three cascade reactions were first optimized separately and possible cross-interactions were determined. Afterwards, the catalytic steps were combined in either a 1-pot-2-steps, 1-pot-3-steps or 2-pot-3-steps cascade to synthesise the desired amino alcohol products on preparative scale.

4.1 Heterologous expression of the enzyme catalysts

The styrene monooxygenase from *Rhodococcus sp.* ST-10 is the first example of a styrene monooxygenase able to perform the epoxidation of aliphatic alkenes, and it has been previously successfully expressed in *E. coli* BL21 (DE3).[74] However, Toda *et al.* reported that co-expression of a chaperon system further improved protein solubility, especially for StyB.[74] Within this work, however, expression in *E. coli* C43 (DE3) has been performed for comparison. *E. coli* C43 (DE3) is a strain optimized for the soluble production of membrane proteins as well as toxic and/or insoluble proteins and could, thus, improve StyAB production without the use of a chaperone system. Using *E. coli* C43 (DE3) already allowed soluble production of the two cytochrome P450s and the respective ferredoxin reductase and ferredoxin[159,181] with moderate to high yields per litre culture (see section 3.1.1.3). However, the expression of StyAB in *E. coli* C43 (DE3) resulted in a twofold lower epoxidation activity compared to the activities obtained for the production of StyAB in *E. coli* BL21 (DE3) (Table 18). In agreement with previous findings,[74] the addition of a chaperone system resulted in a slight improvement in activity compared to the chaperone-free system. The benefits, however, were too minor to justify co-expression of the chaperone system. Moreover, even a low soluble StyB expression should provide enough reduced FAD cofactor to maintain StyA activity at its maximum reaction velocity, as the specific activity of StyB was shown to be several thousand times higher than the specific activity of StyA.[74]

E. coli BL21 (DE3) was also the preferred expression host used for most of the halohydrin dehalogenases with purified protein yields (2-26 mg/L culture) comparable to the ones obtained for the cytochromes P450 enzymes produced in *E. coli* C43 (DE3) (see section 3.1.2.1).[167] With 160 mg/L culture, MsAct gave the highest yield of purified protein compared to other enzymes produced in *E. coli* BL21 (DE3) in this work. Comparing StyAB production in co-expression experiments with alcohol dehydrogenases required for cofactor regeneration, StyAB production was significantly reduced by LkADH co-expression compared to LsADH co-expression. LkADH appears to have a 5-10 fold higher overproduction compared to LsADH, In contrast, co-expression of *styab* with *Lsadh* resulted in a balanced production of both enzymes, in agreement with previous findings (Figure 10).[74] Protein production in *E. coli* depends on the higher demand for transcription or translation caused by the heterologous protein inserted as well as external conditions.[191] Thus, *Lkadh* overexpression may have suppressed StyAB production by occupying cellular machineries such as ribosomes, polymerases, or chaperones, whose abundance may be limiting.[191]

Compared to shake flask experiments, cultivations in a bioreactor generally resulted in a 20% higher final optical density. This is likely due to better aeration during cultivation preventing oxygen limitations compared to shake flask expression. A lack of oxygen triggers the expression of over 200 genes utilized to adapt the cellular metabolism to lower O_2 concentration, hence, resulting in slower cellular growth.[192] Interestingly, in bioreactor cultivations for the production of the HHDHs the final OD_{600} reached was significantly lower compared to StyAB cultivations, although only one antibiotic was supplied. In contrast, cultivation of cells co-expressing StyAB and LsADH required two antibiotics due to the presence of two plasmids. However, both plasmids (pETDuet-*styab* and pCDFDuet-1-*Lsadh*) carry resistance proteins that inactivate the respective antibiotic's molecules by degradation or chemical modification[193,194] In contrast, kanamycin is not degraded by the resistance protein present in pET28 carrying the HHDHs[195], hence more cells harbouring StyAB and LsADH could grow towards the end of cultivation as the actual antibiotic concentration is lower, whereas the antibiotic concentration in HHDHs fermentations is constant throughout the fermentation.

HheE and HheE5, however, generally reached higher expression levels and this might have limited the growth[191] compared to StyAB cultivations. Despite the better growth in the bioreactor compared to shake flasks, the resulting specific activity (U/mg of total protein content) of HheE5 was significantly lower than for shake flask expression. For HheE, this difference was less pronounced. In previous findings, optimal conditions for HheE5 production in *E. coli* BL21 (DE3) were found to be 7 h expression after induction at 30 °C.[167] Possibly, HheE5 folds at a slower rate and during the translation, its

hydrophobic stretches could interact causing the accumulation of inactive enzyme.[192] Better growth in the bioreactor could have led to faster transcription and translation and perhaps more inactive enzyme in agreement with SDS-PAGE gels where HheE5's band is markedly thicker in the insoluble fraction (Appendix, Supporting figure 1).

Similarly, StyAB's specific activity decreased with longer expression times in the bioreactor. In this case, antibiotic degradation may have resulted in an enrichment of cells not carrying the plasmids and lower specific activity. Hence in StyAB production, cells should be harvested roughly 12 hours after induction to maximise optical density and enzyme activity.

4.2 Epoxidation of linear alkenes

Comparing the performance of all eight catalysts screened for epoxidation of substrate **2**, only StyAB and the Shi diketal epoxidation catalyst were able to reach >50% conversion. The remaining catalysts showed less than 5% conversion and poor product enantiomeric excesses (Figure 15), whereas both cytochromes P450 did not show any product formation.

The Jacobsen catalyst, for instance, is designed to perform epoxidation of terminal alkenes and thus, its low activity in the epoxidation of non-terminal alkene **2** is not surprising. However, it is interesting that the catalyst in *S,S*-configuration produced more of the (*R*,*R*)-epoxide and *vice versa*. The unspecific peroxygenase from *Agrocybe aegerita* resulted only in very low conversion which may be, in part, due to non-optimal hydrogen peroxide supplementation. H_2O_2 is needed for catalytic activity and at the same time, it is extremely toxic for the heme prosthetic group present in the enzyme.[63,196] Moreover, the product enantiomeric excess was only close to 50% and hence, not sufficient for implementation in the cascade. Luckily, StyAB and the Shi epoxidation diketal catalyst produced opposite epoxide enantiomers, enabling two different pathways for the synthesis of opposite amino alcohol enantiomers.

4.2.1 Styrene monooxygenase-catalysed epoxidations

The styrene monooxygenase is a two-component system that requires NADH and $FADH_2$ necessary to catalyse the epoxidation; it also requires molecular oxygen for its activity. The catalyst can be applied as whole cells or cell-free extract using cellular metabolism or a cofactor regeneration enzyme to recycle the needed cofactors.

Overall, supplying StyAB as cell-free extract with a cofactor regeneration enzyme always resulted in higher conversion compared to whole-cell reactions. Three factors might be responsible for this: i) oxygen availability, ii) cofactor regeneration strategy and iii) substrate/product inhibition. When using a balloon of pure oxygen instead of O_2 from the air, the conversions in both whole-cell and cell-free extract reactions improved

significantly. The performance of whole cells, however, still remained lower compared to reactions with cell-free extract. In whole-cell reactions, oxygen is being consumed by the cellular metabolism as well (resulting in growth), thus, its availability for the enzyme is lower. The optimal catalyst concentration ensuring sufficient oxygen, with no active input, for StyAB activity despite consumption *via* cell metabolism appeared to be 15 g/L (wet cell weight) and; adding more whole-cell catalyst decreased StyAB's activity further (Figure 18). Subsequently, the different oxygen consumption between whole cells or cell-free extract was quantified using an oxygen sensor. The dissolved oxygen concentration quickly dropped to zero when whole cells were present in solution, whereas in the presence of cell-free extract an equilibrium between O_2 consumed by the epoxidation reaction and O_2 diffusing in the solution was reached (Figure 33, A). The active input of compressed air (3 L/h) resulted in only low improvements in oxygen availability for whole-cell reactions; in contrast, it maintained oxygen saturation in the presence of cell-free proteins (Figure 33, B).

The efficient provision of NADH cofactor by the cellular metabolism strongly depends on the metabolic activity of the cells and the used co-substrate. The alkene substrate used in this work for StyAB may be toxic for the cells resulting in a slower cellular metabolism and hence, slow NADH regeneration in whole-cell reactions. In contrast, a cofactor regeneration enzyme does not require an intact cell and StyAB (as well as the regeneration enzyme) can be supplied as CFE. A cell-free system is independent of substrate toxicity for the metabolism and it also removes mass transfer through the cellular membrane which is normally one of the limiting steps.

Toda *et al.* previously demonstrated that StyAB suffers from both substrate and product inhibition[74] and one strategy to overcome substrate inhibition is by supplementing lower amounts multiple times so that the substrate concentration is kept relatively low but reaching the total amount to be added over time. This strategy allowed to double the achieved epoxide concentration in reactions where LkADH (as separate CFE) was added as cofactor regeneration system (Orange in Figure 32); however, product inhibition was still present. A second organic phase is often used to prevent substrate and product inhibition, as it acts as a reservoir for hydrophobic compounds, such as the ones used in this work, while keeping substrate and product concentrations in the aqueous phase low. The use of hydrophobic solvents (logP>4) in StyAB reactions could improve the overall epoxide concentration. Particularly heptane and octane resulted in relatively high epoxide production, whereas the presence of cyclohexane yielded only moderate improvements. Using more polar organic solvents, such as ethyl acetate, pentyl acetate or hexyl acetate,[68] resulted in no product formation and significant protein degradation after just 2 hours. Toda *et al.* used whole cells that may have protected the

enzyme from the contact with the polar solvents used,[68] although using the same reaction conditions with whole cells but on larger scale did not show any product formation (data not shown). Organic solvents can disrupt the cellular membrane resulting in metabolism damage and cell death,[197] whereas in contact with enzymes they can disturb the network of non-covalent interactions necessary to keep intact the enzyme's secondary and tertiary structures.[198] Interestingly, cyclohexane, pentyl- and hexyl acetate have more similar logP values (3.2, 2.30 and 2.87, respectively), however product formation was observed only when cyclohexane was used as second phase. Very likely, the two acetate esters can disrupt more easily the enzyme's secondary and tertiary structures, whereas cyclohexane with no polar groups has a less dramatic effect. Generally, using organic solvents with a high logP (>4), such as heptane, can limit the interaction between enzymes and organic solvent only at the interface between the two phases. In contrast, using a more polar organic solvent (logP<3), hence more soluble in water, can result in more severe enzyme inactivation compared to less soluble organic solvents. Using heptane or octane as second phase resulted in similar product yields, however heptane has a lower boiling point and would simplify solvent removal during reaction workup. In short, StyAB-catalysed epoxide production could be significantly improved by active air input to remove oxygen limitations. Furthermore, using a cofactor regeneration enzyme and a second organic phase to limit substrate and product inhibition resulted in 100-fold improvement overall.

Additionally, enzyme inactivation appeared to be another factor that seemed to limit the final epoxide concentration reached. On small scale, StyA lost almost all of its activity within the first 6 hours. This may have been caused by mechanical stress due to magnetic stirring as Toda *et al.* performed the epoxidations for 48 hours using horizontal shacking.[68] The set up used in this work for large scale reactions required overhead mechanical stirring that may also cause protein damage due to mechanical stress. Furthermore, the used organic solvent can cause enzyme inactivation as mention above. Hence, a substrate plus enzyme feed was used to further improve epoxide production compared to only substrate feed (Red in Figure 32).

Using optimal conditions for the StyAB-catalysed epoxidation of 50 mM *trans*-2-heptene (**2**), a maximum conversion of 70% (35 mM epoxide) in 80 mL scale could be reached. A further increase in the reaction volume to 250 mL and 100 mM substrate concentration led to a final epoxide concentration of 69 mM (69% conversion). Reaction engineering could increase product formation almost 100 times and the use of a bioreactor, enabling even better control over several variables (*e.g.* dissolved oxygen and substrate feed), may further increase the overall reaction performance. Using a 30 L fed-batch reactor, Panke *et al*. could produce 388 g of styrene oxide in a using growing cells harbouring

the styrene monooxygenase from *Pseudomonas sp.* strain VLB120 and using its cellular metabolism for cofactor regeneration.[76] On the other hand, using a solvent tolerant microorganism as heterologous host could be used to prevent enzyme inactivation caused by the organic solvent as the cell membrane would not be disrupted as easily as the cell membrane of non-tolerant microorganisms. For example, *Kocuria rhizophila* DC2201 has been successfully applied for the production of industrially relevant epoxide concentrations.[199] Another example of a solvent tolerant organism used as heterologous host is *Pseudomonas taiwanensis* VLB120Δ*C* that was successfully applied in a 3.1 L reactor for the production of styrene oxide.[200]

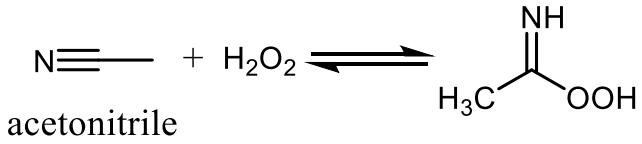
If reaction engineering is not sufficient, protein engineering is often used to improve the enzymes' activity, stability or selectivity. Recently, the styrene monooxygenase from *Pseudomonas sp.* was engineered by fusing its two subunits resulting in a more active catalyst.[201] This styrene monooxygenase fusion results in a 1:1 ratio of the two subunits; thus $FADH_2$ can easily diffuse from StyB to StyA and can be readily used in the epoxidation by StyA minimizing uncoupling.[201] Moreover, StyB solubility was indirectly improved by the production of a single protein containing both subunits. This approach could possibly also improve the catalytic efficiency of the styrene monooxygenase used in this work.

Furthermore, different cofactor regeneration systems might influence the epoxidation reaction differently. For example, cofactor regeneration efficiency using alcohol dehydrogenases is driven by the thermodynamic equilibrium of the regeneration reaction, usually oxidation of an alcohol. As the alcohol is oxidised, the corresponding ketone/aldehyde builds up until the reverse reaction becomes more favourable.[202] This could be a limitation if there is not enough alcohol present to keep up the NADH regeneration for the reaction. At the same time, high alcohol concentrations can negatively affect the enzymes' stability. Alternatively, formate dehydrogenases, for example, oxidise formate to reduce NAD^+. The by-product, CO_2, can freely diffuse out of the reaction mix shifting the reaction equilibrium towards NADH reduction. Moreover, the co-substrate formate has usually no negative or inhibitory effect on enzymes. Testing other cofactor regeneration system could result in further improvements in epoxide production.

4.2.2 Epoxidations performed with the Shi epoxidation diketal catalyst

In contrast to the styrene monooxygenase, the Shi epoxidation diketal catalyst was selected as it produces the (*R*,*R*)-enantiomers of **21** and **22** selectively, however, with a slightly lower enantiomeric excess (>90%, Figure 15). The ee_P in this reaction strongly

depends on the catalyst's stability throughout the reaction. Moreover, the reaction is run in presence of 20-30% acetonitrile at low temperatures and pH > 8, however optimal reaction conditions varied with the type of substrate used.[203] Acetonitrile (logP -0.334) is a water-miscible solvent and could negatively affect the HHDH in the second step of the cascade. Hence, different acetonitrile concentrations were tested in the Shi epoxidation to determine the minimum required amount that would still grant good conversion and high ee_P. Generally, decreasing acetonitrile concentration strongly reduced both conversion and ee_P to around 50% (Figure 43). Likely, the lower acetonitrile concentration might have had a positive effect on the amount of Oxone© dissolved in the aqueous phase. This would result in a higher excess of Oxone© compared to the Shi epoxidation diketal catalyst leading to a faster catalyst degradation via Baeyer-Villiger reaction pathway leading to acetone formation (Scheme 6). This way, acetone could build up faster, and catalysing an unselective epoxidation resulted in a lower final ee_P. Whereas, at higher acetonitrile concentrations the Oxone© solubility is lower causing a slower reaction rate and longer Shi epoxidation diketal catalyst half-life. Furthermore, Shi *et al.* postulated that acetonitrile may participate in the reaction mechanism by mediating the dioxirane formation reacting with the oxidant, H_2O_2 in their case, forming peroxyimidic acid (Scheme 24).[203] Perhaps, this may also be the case and might explain why lowering acetonitrile concentration resulted in lower product yield and ee_P.



Scheme 24. The dioxirane necessary for the Shi epoxidation diketal catalyst may be generated via peroxyimidic acid produced by hydrogen peroxide reacting with acetonitrile (adapted from Shi *et al.*[203])

Another factor to consider here is substrate and catalyst solubility. The Shi epoxidation diketal catalyst, as well as substrates 1 and 2, are hydrophobic compounds with only low solubility in aqueous buffers. Epoxidation reactions run at 25 °C resulted in low ee_P in the conversion of alkene **2** (data not shown) and only with a temperature of -2 °C an ee_P>90% was reached. In the other hand, a temperature of -2 °C would further decrease substrate and Shi epoxidation diketal catalyst solubility. Notably, at low acetonitrile concentrations (5% or 10%) only around 30% conversion was reached, meaning that theoretically, each molecule of catalyst reacted only 1.5 times. It is very likely that only a small fraction of the catalyst was actually dissolved in solution and could react. This small amount of dissolved catalyst can then easily decompose via the Baeyer-Villiger pathway due to the high excess of Oxone© in solution. At this point, acetone, being more soluble

in water than the enantioselective catalyst, can catalyse unselective substrate epoxidation resulting in the poor ee_P measured (Figure 43). In conclusion, acetonitrile concentration has a profound influence on the reaction performance and preferably the used acetonitrile concentration should be above 20% (v/v).

4.3 Epoxide ring opening using HHDHs

Regioselective epoxide ring opening in the 2nd step of the cascade from non-terminal alkenes to amino alcohol products was essential for the successful synthesis of enantiopure amino alcohols. As described previously, HHDHs are typically applied in the ring opening of terminal epoxides displaying high regioselectivity for nucleophilic attack of the terminal carbon. In the azidolysis of non-terminal epoxides, however, the observed regioselectivity was enzyme- and substrate-dependent.

Of the five different tested epoxides, most HHDHs could convert *trans*-2,3-epoxyhexane (**21**) and *trans*-2,3-epoxyheptane (**22**) and fortuitously, seven enzymes displayed extremely high selectivity in the production of azido alcohols **31a** and **32a** with the best being HheA3, HheC and all E-type enzymes.

In contrast, a significantly lower number of HHDHs could convert also more complex epoxides. For example, *trans*-4,5-hexan-1-ol (**23**) carries an additional hydroxyl group at the opposite end of the molecule with respect to the epoxide ring. This additional hydroxyl group suppressed the activity of all HHDHs except HheG. Docking experiments with HheC and HheD2 highlighted that this hydroxyl group may interact preferentially with the catalytic residues resulting in a reversed binding mode (Figure 22 and Figure 23, respectively). This opposite binding mode was not observed for HheG in agreement with experimental data. Co-crystallization of the epoxide in one of the HHDH's active site could further confirm these data. Additionally, kinetic studies of epoxide ring opening in the presence of a terminal alcohol (*e.g.* 1-hexanol) could determine if a compound with a terminal alcohol can act as a competitive inhibitor by interacting with the catalytic triad.

In *trans*-3,4-epoxyheptane (**24**) the epoxide ring is positioned more internally in the molecule compared to epoxide **22.** This reduced the number of HHDHs able to convert **24** compared to epoxide **22**. Furthermore, all enzymes able to convert epoxide **24** showed no regioselectivity in the azidolysis of **21** and **22** (Figure 21). Non-regioselective HHDHs may have a broader active site where epoxides with larger substituents can still adopt the correct geometry needed for catalysis. Epoxide **25** present a benzylic ring substituent instead of an aliphatic chain as for other epoxides tested in this work. Only HheG and HheG2 displayed significant activity towards this epoxide, whereas conversions using HheD3 and HheF were only slightly above the chemical background. Interestingly, HheC showed activity on the styrene oxide but no conversion with **25**. In

this case, the resonance with the benzylic group favours the nucleophilic attack at the more sterically-hindered carbon, however HheC favoured the nucleophilic attack at the less sterically-hindered carbon. In contrast, HheG catalysed the azidolysis of styrene oxide with α-position regioselectivity,[204] as seen for epoxide **25**. In contrast, HheD3 and HheF displayed preference for the formation of regioisomer **35a** even though with very low conversions compared to HheG. For these enzymes, the epoxide ring seems to have adopted the correct geometry with serine and tyrosine to promote nucleophilic attack at the chemically un-favoured β-carbon.

The active site's shape may influence the substrate scope and regioselectivity of the different HHDHs. Comparing the active sites of HHDHs with available crystal structure, it is noticeable that compared to HheD2 and HheG (PDB-ID: 5O30[105]), HheA2 (PDB-ID: 1ZMO[205]), HheB2 (PDB-ID: 4ZD6[206]) and HheC (PDB-ID: 1PWZ[184]) possess a narrow tunnel leading to the active site with the nucleophile binding pocket (Figure 52). Hence, in these cases the converted substrates (**21** and **22** for HheC) could only bind productively with the smaller substituent at the epoxide ring entering first (Figure 22, A and B). In contrast, HheD2's active site appears to extend in both sides of the nucleophile binding pocket (unpublished data), and HheG active site cavity is broader and open compared to all other HHDHs enabling also the conversion of more bulky substrates.[105] This extra space may have allowed the di-substituted epoxides to adopt productive binding having the bulkier substituent on either side of the epoxide ring. This way, in non-regioselective HHDHs, the nucleophilic attack is performed on either carbon of the epoxide ring depending on the geometry with the catalytic triad (Scheme 25).

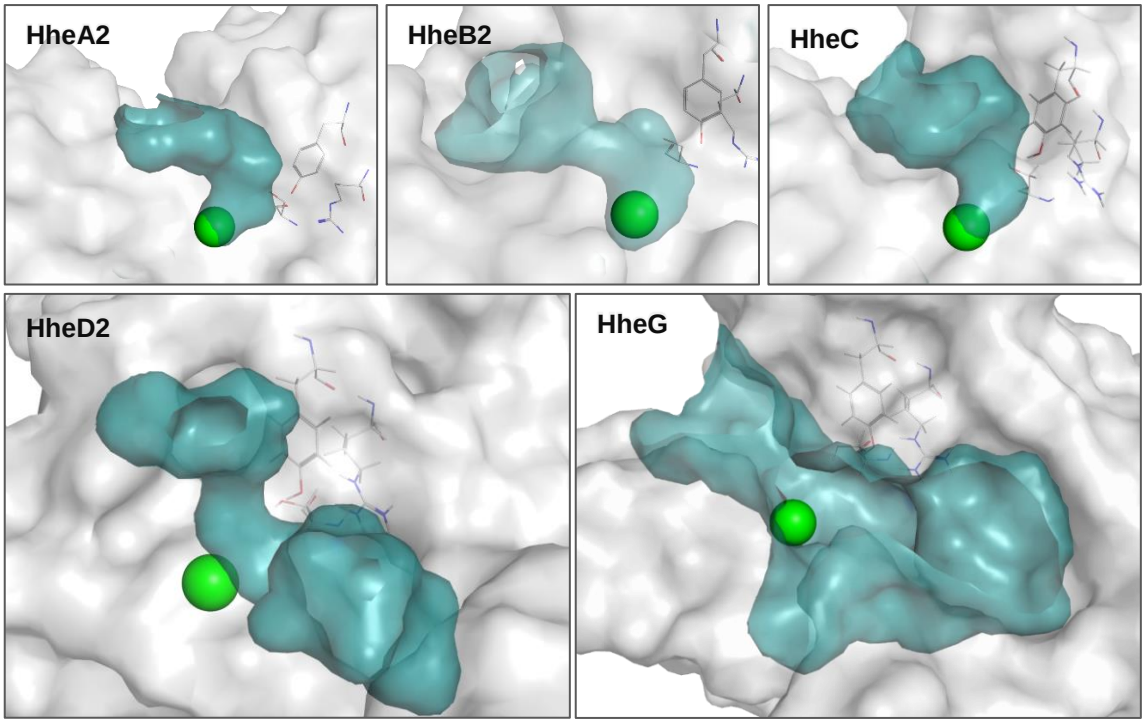
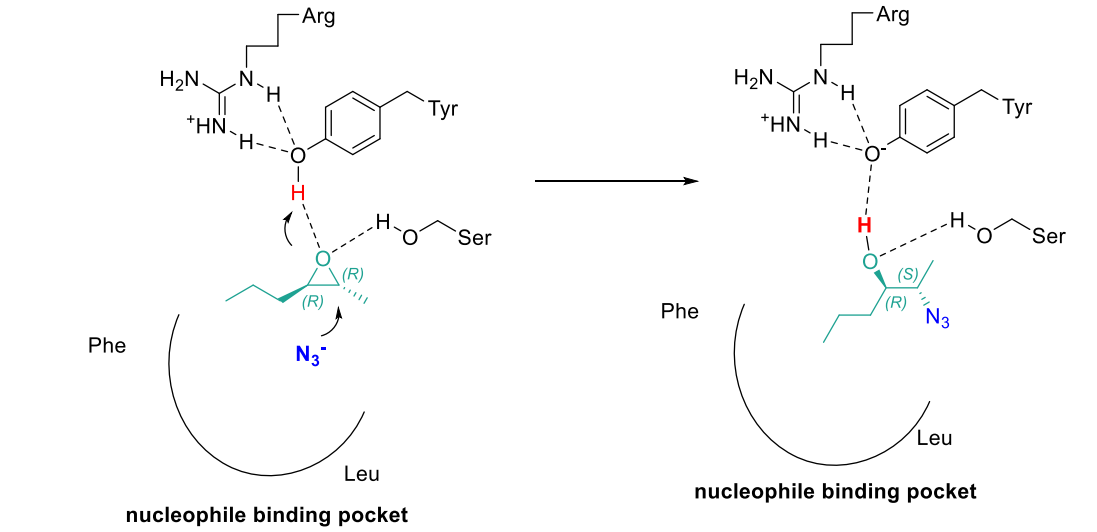


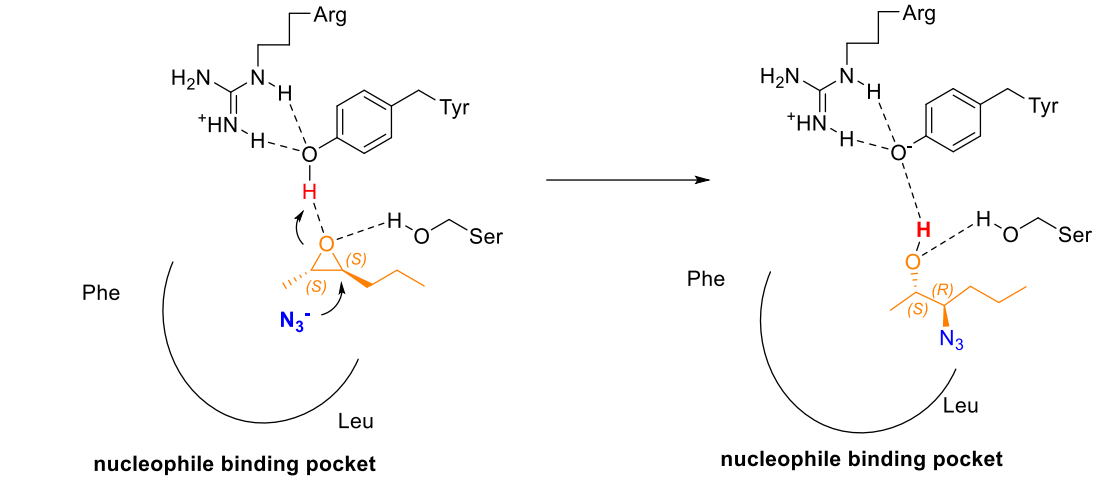
Figure 52. Active sites highlighted in deep teal for HHDHs with available crystal structure [HheA2 (PDB-ID: 1ZMO[205]), HheB2 (PDB-ID: 4ZD6[206]), HheC (PDB-ID: 1PWZ[184]), HheG (PDB-ID: 5O30[105]) and HheD2 (unpublished data)]. The water molecule present in the nucleophile binding pocket is shown as a green sphere to indicate the position of the nucleophile.

Determination of product enantiomeric excesses (ee_P) in the ring opening of racemic epoxides **21** and **22** evidenced that most of HHDH form the corresponding azido alcohols with low to moderate enantioselectivity (Table 21). Interestingly, azido alcohols **31** and **32** were formed mostly in 2*S*,3*R*-configuration (Table 21). Meaning that regioisomers **31a** and **32a** were preferentially formed by nucleophilic attack of the epoxides in (*R,R*)-configuration at the sterically less hindered carbon, whereas regioisomers **31b** and **32b** were preferentially formed by nucleophilic attack of the epoxides in (*S,S*)-configuration at the sterically more hindered carbon. Generally, product enantiomeric excesses (ee_P) for regioisomers **21a** and **31a** were rather low, in contrast formation of regioisomers **31b** and **32b** occurred with higher enantioselectivity in non-regioselective HHDHs. One explanation could be that, in non-regioselective HHDHs, the (*R,R*)-epoxide preferentially enters the catalytic site with the shorter chain first, whereas the (*S,S*)-epoxide preferentially adopts enters with the longer chain first. This way, the geometry between the epoxide ring and catalytic triad favours nucleophilic attack at opposite carbons of the epoxide ring (Scheme 25). Interestingly, HheG and HheD5 produced regioisomers **a** and **b** in equal amounts when racemic epoxides were used as substrates (Table 21). However, they preferentially synthesised regioisomers **31b** (78% for HheG) and **32b** (60% for HheD5) starting from the (*S,S*)-epoxide synthesised in cascade reactions by StyAB (Results, section 3.1.4). The change of regioselectivity in these two enzymes may validate the hypothesis that opposite enantiomers preferentially adopt different productive bindings (Scheme 25). Opposite binding modes for different enantiomers were found in docking studies for the non-regioselective HheG and HheD2, whereas for the regioselective HheC only one binding mode was found for both epoxide enantiomers.

Epoxide ring opening:



A) canonical binding



B) flipped binding

Scheme 25. Possible epoxide ring opening scheme for non-regioselective HHDHs. Opposite epoxide enantiomers have reverse binding. The *R,R*-epoxide may preferentially adopt the canonical binding (A), whereas the *S,S*-epoxide may adopt a flipped binding mode (B) resulting in nucleophilic attack at opposite carbons of the epoxide ring. *Trans*-2,3-epoxyhexane (**1**) is used as model substrate and the enantiomer in *R,R*-configuration is shown in green, whereas the enantiomer in *S,S*-configuration is shown in orange.

4.3.1 Regioselective HHDHs for application in cascades

Among the regioselective HHDHs, HheA3 and HheC are well characterised enzymes already, a deeper investigation was performed on enzymes of the E-type, which are less studied. From the initial screening, HheE and HheE5 appeared to be the best candidates for the conversion of **21** and **22**. A more detailed study of all E-type HHDHs confirmed HheE and HheE5 to be the best enzymes for epoxide ring opening of **21** and **22**, as both displayed significantly better performance than HheE2, HheE3 and HheE4. In general,

HheE showed the best conversion in the azidolysis of **21**, whereas HheE5 was faster in the conversion of **22**. Analysis of the kinetic data confirmed that HheE5 is a significantly better catalyst than HheE in the epoxide ring opening of **22** with a maximal velocity 1.3-fold higher. HheE showed significantly less affinity for epoxide **21** (1.5 timed higher K_M) and lower maximal reaction velocity. However, in whole cell experiments HheE appeared to be much faster than all other enzymes of the E-type and it was selected for cascade reactions starting from substrate **1**. Whereas, HheE5 was selected for cascade reactions starting from substrate **2** having better performance in the conversion of epoxide **22.**

The difference in maximal reaction velocity of HheE might have been influenced by the fixed nucleophile concentration (50 mM). This value is only ~4-fold higher than HheE's K_M for azide whereas it is ~8-fold higher than the K_{50} of HheE5 for azide. As a rule of thumb, a 10-fold higher concentration than the enzyme's K_M/K_{50} should be used to ensure maximal reaction velocity, hence HheE maximal reaction velocity might be less accurate than HheE5.

At higher substrate loadings (50-300 mM), using two equivalents of sodium azide, the spontaneous azidolysis of epoxide **21** was considerably higher than that of epoxide **22** in reactions with HHDH-containing whole cells. This may be related to the higher hydrophobicity of **22** carrying an additional CH_2 group. Thus, epoxide **22** may diffuse more easily through the cellular membrane, where the azide concentration would be lower compared to the rest of the reaction solution. In contrast, a higher fraction of epoxide **21** may be left out of the cellular membrane where chemical azidolysis happened with higher probability due to the high concentration of azide added in the reaction solution. This could be proven by testing if the background reaction is the same for both epoxides in the absence of whole cells. On the other hand, epoxide **21** might be more reactive than **22** resulting in higher chemical background.

The high chemical background reaction could be avoided by decreasing the nucleophile concentration as it is only important to maintain azide concentrations above the enzymes' K_M/K_{50} values to ensure an optimal reaction rate. For substrate concentrations below 10-20 mM, it might be required to add a larger excess of sodium azide, however, at high substrate loadings a ratio of 1:1.1 (substrate:NaN_3) should be sufficient to provide enough nucleophile to reach full conversion.

Preparative scale using only 50 mM substrate were in agreement with small scale reaction and the same high regioselectivity was observed for both enzymes. However, rather low isolated yields were obtained (50% for **31a** synthesised by HheE and 48% of **32a** synthesised by HheE5), this most likely can be improved by better downstream processing as some product loss occurred during product purification.

4.4 Chemical hydrogenation of azido alcohols

The reduction of azido groups to their respective amino groups is a well-known reaction. As mentioned in section 3.1.3, several metals can split H_2 on their surface and this property can be exploited to catalyse the addition of hydrogen to molecules that can accept it. The last step of the synthetic strategy adopted in this work involves hydrogenation of the azido alcohol produced in the first two steps to the corresponding amino alcohol. Using palladium as metal seemed to offer the best compromise between mild reaction conditions and catalyst costs.

Generally, palladium is quite reactive and it essentially reacts with most reducible functionalities. Chemoselectivity can be changed using heterogeneous palladium on different supports or with adequate catalyst poisoning.[120] For instance, entry 2 in Figure 26 (Section 3.1.3) is palladium on $CaCO_3$ with Pb poisoning and it was designed primarily to catalyse the hydrogenation of alkynes to alkenes without further reduction of the resulting alkenes. Therefore, different supports were tested in the hydrogenation of azido alcohol **32** to determine which the best support for hydrogenation of azides.

Using carbon as support resulted in higher activity compared to other matrices. The two slightly different carbon supports with 10 wt.% palladium yielded higher conversion compared to palladium hydroxide on carbon, although it contained a palladium content of 20 wt.%. Here, it seems that palladium in the form of palladium hydroxide has lower activity in the hydrogenation of **32** despite having twice as much of catalyst loading on the carbon support (entry 3, Figure 26). Overall, the palladium (as metal ions) on carbon catalysts (entries 1 and 5, Figure 26) performed equally well. Hence, the selection of the best catalyst for the final cascade was based on catalyst price. The Degussa type E101 carbon support (entry 5, Figure 26) appeared to be 1.5 times more expensive (Sigma-Aldrich quotation) than the regular activated carbon (entry 1, Figure 26). Therefore, the latter was used for hydrogenation of azido alcohols **31** and **32** in the final cascades.

Heterogeneous palladium and nanoparticles in general are known to interact with proteins and living cells causing mutual inactivation as the latter can adsorb on Pd/C surface.[122] Ferritin from *Pyrococcus furiosus* or the commercially available ferritin from horse spleen were tested as a compartmentalization strategy to encapsulate the nanoparticles and overcome the mutual inactivation. This way substrate and hydrogen can freely diffuse through the protein wall isolating the nanoparticles from harmful compounds that reduce the activity of the metal, such as free enzymes or whole cells (Scheme 14). Ferritin's quaternary structure is a 24mer suitable to encapsulate different metal ions. Unfortunately, using the self-produced thermostable ferritin from *P. furiosus* did not show any nanoparticle formation probably due to an incomplete iron removal that prevented palladium encapsulation. On the other hand, using the commercial ferritin from

horse spleen resulted in the formation of the complex ferritin-PdNPs which led to full conversion of 5 mM of **32** within 4 h. However, the synthesis of the complex ferritin-PdNPs requires several labour-intensive steps and an intermediated workup to remove CFE or whole cells before the hydrogenation reaction would be more convenient. Hence, Pd/C was chosen as the candidate catalyst for the last step of the cascade giving the commercial availability and high reactivity.

4.5 Combination of the catalysts

Using a non-terminal *trans*-alkene as a substrate for the synthesis of enantiopure amino alcohol presents several challenges. Most prominent is the control of stereoconfiguration of the chiral centres as well as regioselectivity during functionalization of the molecule as up to four isomers can be generated (Scheme 19). As mentioned previously, the strategy to control the formation of only one amino alcohol enantiomer was to combine an enantioselective epoxidation with a regioselective epoxide-ring opening. The combination of styrene monooxygenase with HheE or HheE5 and Pd/C should lead to the synthesis of (2*R*,3*S*)-2-amino-3-hydroxyhexane (**41a**) and (2*R*,3*S*)-2-amino-3-hydroxyheptane (**42a**) starting from *trans*-2-hexene (**1**) and *trans*-2-heptene (**2**), respectively. In contrast, using the Shi epoxidation diketal catalyst in the epoxidation step should result in the formation of opposite product enantiomers, (2*S*,3*R*)-2-amino-3-hydroxyhexane (**41a**) and (2*S*,3*R*)-2-amino-3-hydroxyheptane (**42a**), respectively. The validity of this strategy was carefully verified by testing different combinations of epoxidation catalyst with regio- or non-regioselective HHDHs (Figure 29 and Figure 30, Section 3.1.4).

4.5.1 StyAB – HHDH – Pd/C cascade

The combination of individual reaction steps, separately optimised before, in a cascade often requires further reaction optimization to achieve high overall productivities. As an example, reaction components required by one catalyst may negatively affect another catalyst in the cascade. Hence, identification of such cross-interactions and -inhibitions is important for cascade setup.

The styrene monooxygenase and the halohydrin dehalogenase work under similar reaction conditions in terms of pH and temperature, and both could be used in the same 50 mM potassium phosphate buffer. This would allow performing both cascade steps simultaneously or in two sequential stages. The presence of azide did not show any effect on the styrene monooxygenase (data not show), however probably at concentrations of azide >100 mM enzyme inactivation could occur as free azide can react with amino acids such as tryptophan cysteine or tyrosine.[207] In contrast, the presence of heptane, required in the epoxidation step to prevent substrate and product

inhibition of StyAB,[74,83] appeared to have a considerable negative effect on the HHDHs (Figure 39). For both enzymes, the conversions were significantly lower with heptane compared to epoxide ring opening reactions without heptane. HheE appeared to be significantly more sensitive to the presence of this organic solvent as only 33% conversion was obtained within 24 hours. HheE may be more sensible to organic solvents and unfold more easily than HheE5. HheE5 was found to have a T_{50} 46 degrees higher than HheE[167] and might explain the lower HheE's stability as thermostability and solvent resistance are often correlated. On the other hand, HheC retained 90% of activity in presence of heptane but its T_{50} is only 9 °C higher than HheE.[167,208]

Secondly, heptane, acting as a sink for hydrophobic compounds, efficiently extracted most of the epoxide from the aqueous phase and; this way, both HHDHs likely could not work at their maximal reaction velocities due to comparably high K_M/K_{50} values and hence low affinity for the substrates (Table 22). Zou *et al*. observed a similar behaviour where higher substrate concentrations in the organic phase resulted in a higher initial rate as more substrate could diffuse in the aqueous phase.[209] Interestingly, epoxide **22** seemed to be more soluble in aqueous phase compared to **21** (Figure 36) hence, a higher starting concentration of **21** in heptane may result in better activity of HheE as more could diffuse in the aqueous phase to react with the HHDH.

Lastly, the influence of palladium on carbon on the first two enzymatic steps was evaluated, although in an ideal industrial process there would be active oxygen input for the first step and hydrogen input in the third step and hence mixing the two gasses (if step 1 and step 3 runs together) is not advisable or strictly forbidden due to potential explosions.

On the other hand, the azidolysis and hydrogenation reactions could in principle work at the same time. According to the experimental data, it seems that the presence of Pd/C had a major negative effect on the epoxidation reaction, especially when StyAB was applied as whole-cell catalyst. Pd/C was shown to be toxic for *E. coli* cell that could not synthesise efficiently the required NADH necessary for the epoxidation;[210] lower conversions compared to control reactions were observed as a result (Figure 40). Epoxidations using StyA as cell-free extract showed a higher product yield and probably only adsorption on Pd/C occurred, but enough reduced cofactor was granted by the presence of BNAH. The epoxide ring opening catalysed by HheE5 seems to be less affected by the presence of the nanoparticles as the cellular adsorption on the Pd nanoparticles may only reduce the available surface for substrate diffusion into the cell.

Afterwards, the influence of reagents of steps 1 and 2 on the hydrogenation reaction was also investigated. Pd/C showed lower tolerance to the presence of whole cells (green line, Figure 41) compared to cell-free extract (orange line, Figure 41). As mentioned

before, enzymes and/or whole cells combined with Pd/C can mutually inactivate each other due to adsorption. The mutual inactivation is concentration dependent and possibly doubling the concentration of cell-free extract would increase the negative effect on the hydrogenation. Nevertheless, based on the cross-interaction between the different catalysts, the best way to run the cascade would be to perform the reaction in three sequential steps in one-pot. The enzymatic catalysts should be supplemented as cell-free extract to minimize their negative influence on the hydrogenation step. The HHDH could be added from the beginning although it is more important to limit the contact with heptane and start the epoxide ring opening reaction after the styrene monooxygenase produces enough epoxide. Finally, the palladium on carbon should be added at the end to complete the cascade.

The combination of the three catalysts in one-pot-three-steps led to a very low isolated yield (2.4%) starting from trans-2-heptene (**2**) as none of the three catalysts could reach full conversion. For this cascade run the main bottleneck seems to be epoxide ring opening step as only 5 mM of 2-azido-3-hydroxyheptane were reached in 48h. Furthermore, only a part of this was further reduced by the Pd/C. Here, the inefficient *in situ* H_2 formation may have contributed to the low yield obtained, Schrittwieser et al. used 10 bar H_2 in an autoclave reactor to reach optimal results in the hydrogenation of azides in cascade with an ADH.[211]

Regioisomeric ratios and ee_P determined for the azido alcohol product were in agreement with previous data and this value should correspond to the amino alcohol product enantiomeric excess, as the hydrogenation step does not involve any racemization.

Besides the low conversion achieved in all steps, an inefficient workup might have further lowered the isolated yield as most of the amino alcohol could not be efficiently extracted from the aqueous phase. To overcome this limitation, the hydrogenation reaction was separated from the first two steps and performed in methanol. This way, free proteins from the first two steps could be removed in the intermediate work up and the obtained crude azido alcohols could be dissolved in an organic solvent (methanol in this case), which led to improved work up efficiency. Separating the last step from the first two resulted in a fourfold increase in isolated yields with no negative influence on the product enantiomeric excess, but slightly lower regioisomeric ratio for HheE-catalysed reactions. This is possibly explained by the presence of heptane resulting in reduced HheE activity. Hence, the impact of the unselective chemical azidolysis is higher resulting in a slight increase in regioisomer **b** formation.

Apart from the improvements discussed for the StyAB-catalysed epoxidation step (section 4.2.1), different strategies could be implemented to further improve the final

product yields. The low solvent resistance of the HHDHs appears to be the limiting factor. Heavy enzyme engineering (29 mutations) could significantly improve the solvent resistance of HheC which was active even at 50% acetonitrile.[212] Alternatively, crosslinking HheC using the CLEA technology resulted only in a 2% loss of activity in presence of heptane.[208] Additionally, physical separation of the catalysts through PDMS was used to prevent enzyme inactivation caused by Cu ions. Sato *et al.* demonstrated that the PDMS membrane acted as a barrier between the chemical and enzymatic catalyst whereas the substrate and product could freely diffuse through it.[213]

4.5.2 Shi epoxidation diketal catalyst – HHDH – Pd/C cascade

The reaction conditions required by the Shi epoxidation diketal catalyst are extremely different from the conditions required in the HHDH-catalysed epoxide ring opening. Hence, both steps in a cascade can only be performed sequentially. As for cascade 1, cross-interactions and –inhibitions between reagents and the catalysts were determined prior to the combination of reaction steps.

As observed for heptane in cascade 1, acetonitrile had a strong negative effect on the halohydrin dehalogenases as well. Overall, both HHDHs were significantly more sensitive to acetonitrile compared to heptane (logP 4.5), probably due to its higher solubility in water and higher polarity (logP -0334). Both enzymes seem to lose activity completely within four hours, as no further conversion was observed after this time point. In a previous study, HheC showed a residual activity around 30% after incubation with 12.5% acetonitrile for 4 h[212] whereas incubation with more apolar organic solvents such as heptane or hexane showed less inactivation,[208] in agreement with the findings in this work. The HHDHs' sensitivity to organic solvents indirectly interfered with the epoxidation step as lower acetonitrile concentrations (5%) led to the formation of only low epoxide concentrations with also low product enantiomeric excess (section 3.2.1).

An Oxone© concentration of 10 mM did not affect the epoxide ring opening reaction, whereas 50 mM resulted in lower final conversion (green line, Figure 45, B). This is very likely an indirect effect due to Oxone© rapidly reacting with N_3 generating N_2 and N_2O explaining the observed intense foaming.[188] As a result, only a limited amount of nucleophile was actually available for the azidolysis reaction. Therefore, it was important to decompose all residual Oxone© from the epoxidation reaction before the addition of azide as nucleophile for epoxide ring opening in the second step.

The attempt to run cascade 2 as a sequential one-pot-three-steps cascade resulted in a slightly higher isolated yield of the final amino alcohol compared to cascade 1 (3% vs 2.4%). The product enantiomeric excess, however, was significantly worse (ee_P = 34%) due to the low acetonitrile concentration used in the first step. This time, the hydrogenation catalyst seemed to have been active for a longer reaction time, and

compared to cascade 1 there is only one enzymatic step. In cascade 1, the final total protein content reached was about 18 g/L which is four times higher than the amount used in the cross-interaction study. Whereas in cascade 2, the same amount of cells (60 g/L) was used as in the cross-interaction study, hence Pd/C inactivation may have occurred at a slower rate.

Again, separating the hydrogenation step from the first two cascade steps (= 2-pots-3-steps cascade) could improve the isolated yields significantly. In general, the isolated yields were roughly two-fold higher than the ones in cascade 1, however poor ee_P (40-50%) was observed in the Shi epoxidation. These values could be slightly improved in the epoxide ring opening step, reaching final ee_P of around 60%. Regioselective HHDHs, like HheE and HheE5 may be stereoselective for epoxides in (*R*,*R*)-conformation as observed for HheC in the epoxide ring opening of styrene oxide,[185] whereas more of the epoxide in (*S*,*S*)-conformation was opened unselectively by the chemical reaction with higher probability. The observed regioisomeric ratio of formed azido alcohols was slightly lower compared to the one measured in the HheE and HheE5-catalysed epoxide ring opening of racemic epoxides *trans*-2,3-epoxyhexane (**21**) and *trans*-2,3-epoxyheptane (**22**), respectively. This may be due to the presence of some residual Oxone© which was shown to catalyse epoxide ring opening in presence of acetonitrile and azide.[214] Luckily, the regioisomers could be separated by column chromatography; however, this resulted in approximately 12% product loss. In conclusion, the combination of Shi epoxidation, HHDH-catalysed epoxide ring opening and chemical hydrogenation of the formed azido alcohols resulted in higher yields, however, at the expense of selectivity. The latter is mainly caused by the low amounts (5%) of acetonitrile used in the Shi epoxidation to ensure compatibility with the subsequent HHDH-catalysed step. *In situ* removal of epoxide product (ISPR) or acetonitrile prior starting the epoxide ring opening reaction would be beneficial to improve isolated yields as the HHDHs would not be inactivated and ee_P as the Shi diketal catalyst could produce the (*R,R*)-epoxide with 93% ee_P in presence of >20% acetonitrile. A detailed study of the major differences between acetonitrile and epoxide products would determine the optimal ISPR strategy to use, example, a specific membrane where either acetonitrile or the epoxide (**21** or **22**) would adsorb allowing an easy separation.[215] Alternatively, the HHDHs could be engineered to improve their solvent tolerance towards acetonitrile, as mentioned before, HheC after protein engineering (27 mutations) could retain its activity after incubation for 4h in 50% acetonitrile.[212]

4.6 Cascade starting from cyclic alkenes

4.6.1 MsAct-catalysed epoxidation of cyclohexene

Unlike for cascade 1 and 2, the strategy for this cascade was slightly different as epoxidation of cyclohexene leads to an achiral epoxide. Hence, cascade 3 would only require a stereoselective epoxide ring opening whereas the epoxidation can be unselective. Lipases, like CAL-B, are known to mediate the epoxidation of alkenes by forming a peracid via perhydrolysis of an acyl donor such as ethyl acetate.[216] The peracid can then perform a chemical epoxidation of the alkene substrate. Generally, this is performed in organic solvents as CAL-B's perhydrolysis/hydrolysis ratio is low and the presence of water results mostly in the hydrolysis of the acyl donor and no epoxidation.[216]

MsAct, in contrast, has a high perhydrolysis/hydrolysis ratio, thanks to its highly hydrophobic active site.[161] Hence, an enzyme that can mediate the epoxidation in a buffer solution would be desirable giving the low solvent tolerance shown by HHDHs in the previous sections.

Overall, the maximum conversion reached in the epoxidation of cyclohexene performed by MsAct was 50% and concentrations of ethyl acetate, as acyl donor, higher than 10% significantly hampered the product yield. Possibly an acyl donor with higher logP such as octyl acetate (logP 3.2) could limit protein damage caused by ethyl acetate (logP 0.73). MsAct optimal pH was shown to be around 8,[161] however acetic acid is being formed during the reaction as a result of the peracetic acid reacting and by ethyl acetate hydrolysis to some minor extent. A shift of the pH during the reaction could have resulted in non-optimal enzyme activity and low product yield. Shifting to saturated $NaHCO_3$ buffer at pH 8, which has approximatively 1M buffering capacity, yielded no product formation. In these reactions, however, high CO_2 formation was observed as a result of the buffering activity. Cycloalkenes and derivatives (*e.g.* cyclohexene and cyclohexene oxide) are known volatile compounds[217] and CO_2 bubbling out might have caused intense product and substrate evaporation as none could be detected in GC analysis. Running the reaction with tight pH control would likely improve the conversion as the optimal pH for the enzyme can be maintained throughout the course of the reaction.

4.6.2 HheG engineering using the Rosetta software suite

There are several ways to improve proteins towards a certain desired characteristic.[5,126] Random approaches (*e.g.* directed evolution) are still powerful tools in protein engineering to identify mutations improving enzyme stability, activity or selectivity if a high throughput assay is available.[5]

It is not rare that a mutation far from the active site could improve either the stability or the activity of a certain enzyme.[218–220] For dihydrofolate reductase (DHFR) from *E. coli*, for instance, several mutations located away from the active site influenced catalytic properties and the interaction with the ligand.[218] Unfortunately, current algorithms cannot predict such second shell mutations yet. However, when the reaction mechanism is elucidated, a rational approach to improve enzyme activity or selectivity may work better than directed evolution as only a restricted number of mutants is generated that need to be tested.[136]

Before this work, it was known that HheG preferentially produces (1*R*,2*S*)-2-cyanocyclohexanol and (1*S*,2*S*)-2-azidocyclohexanol in the epoxide ring opening of cyclohexene oxide using cyanide or azide, respectively. Both reactions resulted in a product enantiomeric excess of 40% and the hypothesis was that cyclohexene oxide could bind in two different orientations with different probabilities in the active site of HheG resulting in the observed ee_P.[105] Hence, stabilising one of the two orientations should result in an increased product enantiomeric excess. Consequently, substrate-bound HheG structures were produced using HheC's substrate-bound crystal structures to correctly place the epoxide ring or the cyano- and hydroxyl group with superimposition. Using HheC's substrate-bound crystal structure was necessary to superimpose the catalytic triad and nucleophile binding pocket residues of the two structures as a substrate-bound structure of HheG is not available. A ligand-bound structure is needed as input for the coupled moves algorithm of Rosetta to select residues around the ligand that would improve protein-ligand interactions.

The coupled moves results highlighted that only a few positions, showing high variability, might be relevant to stabilise the ligand in the starting position. This means that mutations at these positions often had a better energy score than the wild type residue and therefore were accepted more often. For other positions, the wild type residue showed 100% frequency meaning that no other amino acid was accepted during *in silico* mutagenesis. Most likely, these positions do not interact with the ligand or no other amino acid had a better score than the wild-type residue.

As described before, most of the mutants did not improve activity or selectivity and in some cases, a single mutation deleted the activity completely (Figure 50, A). Overall, what emerged was that one position had a significant influence on both activity and selectivity. Most of the mutations at this position had a considerably lower activity and most importantly, the enantioselectivity was reversed in few cases compared to the wild type. Surprisingly, different enantioselectivities were found using cyanide as a nucleophile. Possibly, the mutation at this position modified the nucleophile binding pocket resulting in reverse enantioselectivity in the synthesis of the azido alcohol. At the

same time, it may have had a negative effect on cyanide binding causing lower activity and selectivity. According to the elucidated epoxide ring opening mechanism, a preferential synthesis of (1*R*,2*R*)-2-azidocyclohexanol should correspond to a preferential synthesis of (1*R*,2*S*)-2-cyanocyclohexanol having the same stereoconfiguration. In contrast, all the active mutants showed equal or lower enantiomeric excess compared to the wild type in the cyanolysis of cyclohexene oxide (Figure 51). For HheC, various engineering experiments highlighted that residues in the external part of the catalytic tunnel are responsible for the enzyme's activity and stereoselectivity.[106,107,206] HheG's active site, however, may undergo major rearrangements upon substrate binding.[105] Thus, it is difficult to speculate on the reasons for the observed different product enantiomeric excesses obtained in reactions with different nucleophiles.

To conclude, it was possible to correctly identify a few mutations with a strong influence on enantioselectivity and activity using the coupled moves application in Rosetta. It could be interesting, however, to dock the ligands in opposite poses to see if different mutants would be suggested by the coupled moves. Furthermore, site-saturation mutagenesis of the position that showed the highest influence on HheG's enantioselectivity could result in the identification of additional mutations to further improve the selectivity.

5. Conclusions and future prospects

The aim of this work was to set up a novel synthetic pathway for the synthesis of enantiopure 2,3-amino alcohols through the combination of several chemical and enzymatic steps. Amino- and azido alcohols are important compounds and they can be used either as building blocks or as chiral auxiliaries in the synthesis of complex molecules. Our strategy for the synthesis of enantiopure azido alcohols from non-terminal *trans*-alkenes was to combine a stereoselective epoxidation with a regioselective epoxide ring opening that would allow control of the stereoconfiguration of the two newly formed chiral centres. Finally, palladium nanoparticles were used to catalyse the hydrogenation of the azido group to an amino group without changing the stereoconfiguration.

The most interesting achievement in this work, thus, was the setup and development of the synthetic strategy postulated starting from catalyst identification leading to preparative scale reactions for the production of enantiopure amino alcohols.
Here, the asymmetric epoxidation catalysed by the styrene monooxygenase or the Shi epoxidation diketal catalyst is the key step to control the chiral centres in the 2,3-amino alcohol product, whereas the regioselective epoxide ring opening catalysed by the HHDH is essential to determine which of the two chiral centre will be reversed following the due to the S_N2-type mechanism. The prepared enantiopure amino alcohols are relevant molecules often used as chiral building blocks in the pharmaceutical industry, for example in the synthesis of anti-inflammatory, antiviral, or antitumour agents.
This approach compares favourably with existing synthetic routes for the production of enantiopure amino alcohols, as the number of steps required in the synthesis could be reduced compared to known methods (for examples see Section 1.2). Using a one-pot concept, however, only low isolated yields (2-3%) could be achieved, whereas an intermediate work up before the hydrogenation step could significantly improve isolated yields. The combination of the styrene monooxygenase, HHDH and Pd/C resulted in the production of (*2R,3S*)-2-amino/azido-3-hydroxyhexane and (*2R,3S*)-2-amino/azido-3-hydroxyheptane with an ee_P>99%. In contrast, the cascade combining the Shi epoxidation diketal catalyst, HHDH and Pd/C afforded (*2S,3R*)-2-amino/azido-3-hydroxyhexane and (*2S,3R*)-2-amino/azido-3-hydroxyheptane only with a low ee_P (40-50%), due to incompatibility of the HHDHs with acetonitrile. Overall, the main challenge identified in combining a chemical and enzyme catalysts in this cascade was the incompatibility of reaction conditions. In particular, the presence of acetonitrile from the epoxidation step significantly hampered the epoxide ring opening reaction due to inactivation of the HHDHs. Decreasing the acetonitrile concentration resulted in low

yields in the Shi epoxidation reaction as the epoxide was obtained with only low ee_P. The most promising approach to improve the overall reaction yields would be the laboratory evolution of HheE and HheE5 towards higher solvent resistance. Alternatively, immobilization of the HHDHs might also improve epoxide ring opening efficiency in the presence of ≥20% acetonitrile concentration. Additionally, the strategy chosen in this cascade reactions is highly versatile as changing epoxidation catalysts, HHDH's type and/or nucleophile used in the epoxide ring opening could allow the conversion of a broader array of alkenes for the synthesis of different enantiopure β-substituted alcohols.

The styrene monooxygenase from *Rhodococcus sp.* ST-10 is the first example of a styrene monooxygenase with significant activity on aliphatic alkenes.[85] In this work, the epoxidation of *trans*-2-heptene could be scaled up to a volume of 200 mL reaching 69 mM epoxide product (69% conversion) after careful optimisation of reaction conditions. Hence, the work performed in this project highlights the importance of correctly identifying all parameters with major influence on the epoxidation yield (*e.g.* cofactor regeneration and oxygen availability, substrate and product inhibition) and their careful optimisation for successful upscaling of this monooxygenase-catalysed reaction. Additional reaction engineering would further increase the final epoxide concentration. For example, performing a reaction in a bioreactor allows to precisely control several parameters (*e.g.* dissolved oxygen, substrate and enzyme feed) and it would likely increase product yields and reproducibility.

As mentioned previously, the epoxidation step is a key reaction step to control the chirality of one of the stereocenters in the amino-alcohols produced. In the Shi epoxidation, the opposite epoxide enantiomer is obtained compared to the styrene monooxygenase-catalysed reaction facilitating the formation of opposite amino alcohol enantiomers. Identifying a complementary catalyst to selectively access opposite product enantiomers confirmed the potential and versatility of this new synthetic route. A possible alternative solution could be to engineer or discover a styrene monooxygenase to synthesis the epoxide in (*R*,*R*)-conformation as so far, all known styrene monooxygenases show selectivity towards the opposite epoxide enantiomer.

Another important aspect of this thesis is the detailed investigation of the azidolysis of five vicinally di-substituted epoxides catalysed by HHDHs. For the first time, this study highlighted the ability of HHDHs to accept also non-terminal epoxides (*trans*-2,3-epoxyhexane, *trans*-2,3-epoxyheptane, *trans*-2,3-epoxyhexan-1-ol, trans-3,4-epoxyheptane and trans-β-methyl styrene oxide) as substrates. Before this work, it was state of the art that HHDHs convert only terminal epoxides with almost exclusive nucleophilic attack on the terminal carbon of the epoxide ring. HheG from *Ilumatobacter*

coccineus displayed the broadest substrate scope as it was the only HHDH able to convert all five substrates with good to high activity. Interestingly, many other tested HHDHs converted at least the methyl-substituted epoxides *trans*-2,3-epoxyhexane and *trans*-2,3-epoxyheptane with different regio- and stereoselectivity.
Furthermore, docking studies based on an epoxide-bound structure of HheC as well as on substrate-free structures of HheD2 and HheG revealed first structural insights into the enzymes' observed substrate specificity and regioselectivity. The shape and size of the active site may play an important role in the enzymes' selectivity and substrate specificity. HHDHs with a broad active site such as HheG displayed low regioselectivity and broad substrate scope, whereas enzymes with a narrow tunnel shape-like active site were highly regioselective but fewer substrates were accepted. This further broadened the biocatalytic applicability of HHDHs, in particular HheG, as also sterically more demanding epoxides or vicinally disubstituted halo-alcohols can be used as substrate. Perhaps, this would lead to an even wider exploration of the HHDH's substrate scope by testing epoxides and halo-alcohols with more or less demanding substituents.

Lastly, using an *in silico* assisted rational design allowed the successful identification of HheG variants with increased or inverted enantioselectivity in the epoxide ring opening of cyclohexene oxide using azide as nucleophile. This method proved to be a viable option to identify important residues, especially when available protein engineering information from other HHDHs could only be partially transferred. Furthermore, HheG's stereoselectivity was found to be dependent on the used nucleophile as HheG variants with improved selectivity in the azidolysis reaction displayed no selectivity in the cyanolysis of cyclohexene oxide. A more in-depth study on other HHDHs might give additional insights on HHDHs regio- and enantioselectivity. Moreover, starting from a substrate-bound crystal structure of HheG for rational protein design could yield more accurate predictions and hence, might lead to the identification of even more selective variants.

6. References

[1] R. Singh, M. Kumar, A. Mittal, P. K. Mehta, *3 Biotech* **2016**, *6*, 174.
[2] B. Horowitz, *Sci. Mon.* **1918**, *6*, 253–259.
[3] R. A. Sheldon, P. C. Pereira, *Chem. Soc. Rev.* **2017**, *46*, 2678–2691.
[4] R. DiCosimo, J. McAuliffe, A. J. Poulose, G. Bohlmann, *Chem. Soc. Rev.* **2013**, *42*, 6437–6474.
[5] U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185–194.
[6] U. T. Bornscheuer, *Phil Trans R Soc A* **2018**, *376*, 20170063.
[7] F. H. Arnold, *Chem. Eng. Sci.* **1996**, *51*, 5091–5102.
[8] A. Zaks, A. M. Klibanov, *Science* **1984**, *224*, 1249–1251.
[9] J.-M. Choi, S.-S. Han, H.-S. Kim, *Biotechnol. Adv.* **2015**, *33*, 1443–1454.
[10] D. J. C. Constable, P. J. Dunn, J. D. Hayler, G. R. Humphrey, J. Johnnie L. Leazer, R. J. Linderman, K. Lorenz, J. Manley, B. A. Pearlman, A. Wells, et al., *Green Chem.* **2007**, *9*, 411–420.
[11] A. P. Green, N. J. Turner, *Perspect. Sci.* **2016**, *9*, 42–48.
[12] N. J. Turner, E. O'Reilly, *Nat. Chem. Biol.* **2013**, *9*, 285–288.
[13] A. Bruggink, R. Schoevaart, T. Kieboom, *Org. Process Res. Dev.* **2003**, *7*, 622–640.
[14] F. Rudroff, M. D. Mihovilovic, H. Gröger, R. Snajdrova, H. Iding, U. T. Bornscheuer, *Nat. Catal.* **2018**, *1*, 12.
[15] E. Ricca, B. Brucher, J. H. Schrittwieser, *Adv. Synth. Catal.* **2011**, *353*, 2239–2262.
[16] M. Makkee, A. P. G. Kieboom, H. V. Bekkum, J. A. Roels, *J. Chem. Soc. Chem. Commun.* **1980**, *0*, 930–931.
[17] J. V. Allen, J. M. J. Williams, *Tetrahedron Lett.* **1996**, *37*, 1859–1862.
[18] P. M. Dinh, J. A. Howarth, A. R. Hudnott, J. M. J. Williams, W. Harris, *Tetrahedron Lett.* **1996**, *37*, 7623–7626.
[19] E. Burda, W. Hummel, H. Gröger, *Angew. Chem. Int. Ed.* **2008**, *47*, 9551–9554.
[20] E. F. V. Scriven, K. Turnbull, *Chem. Rev.* **1988**, *88*, 297–368.
[21] F. Fringuelli, O. Piermatti, F. Pizzo, L. Vaccaro, *J. Org. Chem.* **1999**, *64*, 6094–6096.
[22] E. N. Jacobsen, *Acc. Chem. Res.* **2000**, *33*, 421–431.
[23] D. E. White, P. M. Tadross, Z. Lu, E. N. Jacobsen, *Tetrahedron* **2014**, *70*, 4165–4180.
[24] J. E. T. van Hylckama Vlieg, L. Tang, J. H. Lutje Spelberg, T. Smilda, G. J. Poelarends, T. Bosma, A. E. J. van Merode, M. W. Fraaije, D. B. Janssen, *J. Bacteriol.* **2001**, *183*, 5058–5066.
[25] A. Archelas, R. Furstoss, *Curr. Opin. Chem. Biol.* **2001**, *5*, 112–119.
[26] C. E. Castro, E. W. Bartnicki, *Biochemistry* **1968**, *7*, 3213–3218.
[27] H. M. S. Assis, A. T. Bull, D. J. Hardman, *Enzyme Microb. Technol.* **1998**, *22*, 545–551.
[28] O. K. Karjalainen, A. M. P. Koskinen, *Org. Biomol. Chem.* **2012**, *10*, 4311–4326.
[29] S.-L. Shi, Z. L. Wong, S. L. Buchwald, *Nature* **2016**, *532*, 353–356.
[30] D. J. Ager, I. Prakash, D. R. Schaad, *Chem. Rev.* **1996**, *96*, 835–876.
[31] S. Wu, Y. Zhou, T. Wang, H.-P. Too, D. I. C. Wang, Z. Li, *Nat. Commun.* **2016**, *7*, 11917.
[32] J. H. Schrittwieser, F. Coccia, S. Kara, B. Grischek, W. Kroutil, N. d'Alessandro, F. Hollmann, *Green Chem.* **2013**, *15*, 3318–3331.
[33] C. U. Ingram, M. Bommer, M. E. B. Smith, P. A. Dalby, J. M. Ward, H. C. Hailes, G. J. Lye, *Biotechnol. Bioeng.* **2007**, *96*, 559–569.
[34] N. J. Turner, *Curr. Opin. Biotechnol.* **2000**, *11*, 527–531.

[35] T. Katsuki, K. B. Sharpless, *J. Am. Chem. Soc.* **1980**, *102*, 5974–5976.
[36] Y. Hoshino, H. Yamamoto, *J. Am. Chem. Soc.* **2000**, *122*, 10452–10453.
[37] W. Zhang, H. Yamamoto, *J. Am. Chem. Soc.* **2007**, *129*, 286–287.
[38] Q.-H. Xia, H.-Q. Ge, C.-P. Ye, Z.-M. Liu, K.-X. Su, *Chem. Rev.* **2005**, *105*, 1603–1662.
[39] C. Wang, H. Yamamoto, *Chem. – Asian J.* **2015**, *10*, 2056–2068.
[40] W. Zhang, J. L. Loebach, S. R. Wilson, E. N. Jacobsen, *J. Am. Chem. Soc.* **1990**, *112*, 2801–2803.
[41] R. Irie, K. Noda, Y. Ito, N. Matsumoto, T. Katsuki, *Tetrahedron Lett.* **1990**, *31*, 7345–7348.
[42] F. Minutolo, D. Pini, A. Petri, P. Salvadori, *Tetrahedron Asymmetry* **1996**, *7*, 2293–2302.
[43] T. Yamada, T. Takai, O. Rhode, T. Mukaiyama, *Chem. Lett.* **1991**, *20*, 1–4.
[44] Y. N. Ito, T. Katsuki, *Bull. Chem. Soc. Jpn.* **1999**, *72*, 603–619.
[45] T. Miyazaki, T. Katsuki, *Synlett* **2003**, *2003*, 1046–1048.
[46] B. D. Brandes, E. N. Jacobsen, *J. Org. Chem.* **1994**, *59*, 4378–4380.
[47] R. Curci, M. Fiorentino, M. R. Serio, *J. Chem. Soc. Chem. Commun.* **1984**, *0*, 155–156.
[48] Y. Tu, Z.-X. Wang, Y. Shi, *J. Am. Chem. Soc.* **1996**, *118*, 9806–9807.
[49] L. Shu, Y. Shi, *J. Org. Chem.* **2000**, *65*, 8807–8810.
[50] X.-Y. Wu, X. She, Y. Shi, *J. Am. Chem. Soc.* **2002**, *124*, 8792–8793.
[51] S. Wu, Y. Zhou, Z. Li, *Chem. Commun.* **2019**, DOI 10.1039/C8CC07828A.
[52] Y. Liang, J. Wei, X. Qiu, N. Jiao, *Chem. Rev.* **2018**, *118*, 4912–4945.
[53] F. Hollmann, I. W. C. E. Arends, K. Buehler, A. Schallmey, B. Bühler, *Green Chem.* **2011**, *13*, 226–265.
[54] D. Holtmann, M. W. Fraaije, I. W. C. E. Arends, D. J. Opperman, F. Hollmann, *Chem. Commun.* **2014**, *50*, 13180–13200.
[55] E. T. Farinas, M. Alcalde, F. Arnold, *Tetrahedron* **2004**, *60*, 525–528.
[56] D. E. Torres Pazmiño, M. Winkler, A. Glieder, M. W. Fraaije, *J. Biotechnol.* **2010**, *146*, 9–24.
[57] V. B. Urlacher, S. Eiben, *Trends Biotechnol.* **2006**, *24*, 324–330.
[58] F. P. Guengerich, *Nat. Rev. Drug Discov.* **2002**, *1*, 359–366.
[59] J. O. Miners, *Clin. Exp. Pharmacol. Physiol.* **2002**, *29*, 1040–1044.
[60] J. B. van Beilen, W. A. Duetz, A. Schmid, B. Witholt, *Trends Biotechnol.* **2003**, *21*, 170–177.
[61] F. van Rantwijk, R. A. Sheldon, *Curr. Opin. Biotechnol.* **2000**, *11*, 554–564.
[62] E. Churakova, M. Kluge, R. Ullrich, I. Arends, M. Hofrichter, F. Hollmann, *Angew. Chem. Int. Ed.* **2011**, *50*, 10716–10719.
[63] Y. Wang, D. Lan, R. Durrani, F. Hollmann, *Curr. Opin. Chem. Biol.* **2017**, *37*, 1–9.
[64] K. Piontek, E. Strittmatter, R. Ullrich, G. Gröbe, M. J. Pecyna, M. Kluge, K. Scheibner, M. Hofrichter, D. A. Plattner, *J. Biol. Chem.* **2013**, jbc.M113.514521.
[65] W. J. H. van Berkel, N. M. Kamerbeek, M. W. Fraaije, *J. Biotechnol.* **2006**, *124*, 670–689.
[66] B. Entsch, W. J. van Berkel, *FASEB J.* **1995**, *9*, 476–483.
[67] M. Bučko, P. Gemeiner, A. Schenkmayerová, T. Krajčovič, F. Rudroff, M. D. Mihovilovič, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 6585–6599.
[68] H. Toda, R. Imae, N. Itoh, *Adv. Synth. Catal.* **2014**, *356*, 3443–3450.
[69] S. Panke, V. de Lorenzo, A. Kaiser, B. Witholt, M. G. Wubbolts, *Appl. Environ. Microbiol.* **1999**, *65*, 5619–5623.
[70] D. M. Ziegler, *Drug Metab. Rev.* **2002**, *34*, 503–511.

[71] I. Hilker, R. Wohlgemuth, V. Alphand, R. Furstoss, *Biotechnol. Bioeng.* **2005**, *92*, 702–710.
[72] C. V. F. Baldwin, R. Wohlgemuth, J. M. Woodley, *Org. Process Res. Dev.* **2008**, *12*, 660–665.
[73] I. Hilker, M. C. Gutiérrez, R. Furstoss, J. Ward, R. Wohlgemuth, V. Alphand, *Nat. Protoc.* **2008**, *3*, 546–554.
[74] H. Toda, R. Imae, T. Komio, N. Itoh, *Appl. Microbiol. Biotechnol.* **2012**, *96*, 407–418.
[75] M. K. Julsing, D. Kuhn, A. Schmid, B. Bühler, *Biotechnol. Bioeng.* **2012**, *109*, 1109–1119.
[76] S. Panke, M. Held, M. G. Wubbolts, B. Witholt, A. Schmid, *Biotechnol. Bioeng.* **2002**, *80*, 33–41.
[77] R. Gross, K. Buehler, A. Schmid, *Biotechnol. Bioeng.* **2013**, *110*, 424–436.
[78] S. Panke, B. Witholt, A. Schmid, M. G. Wubbolts, *Appl. Environ. Microbiol.* **1998**, *64*, 2032–43.
[79] D. Tischler, M. Schlömann, W. J. H. van Berkel, G. T. Gassner, *FEBS Lett.* **2013**, *587*, 3848–3852.
[80] A. Riedel, T. Heine, A. H. Westphal, C. Conrad, P. Rathsack, W. J. H. van Berkel, D. Tischler, *AMB Express* **2015**, *5*, 30.
[81] M. Oelschlägel, J. Zimmerling, M. Schlömann, D. Tischler, *Microbiology* **2014**, *160*, 2481–2491.
[82] D. Tischler, D. Eulberg, S. Lakner, S. R. Kaschabek, W. J. H. van Berkel, M. Schlömann, *J. Bacteriol.* **2009**, *191*, 4996–5009.
[83] H. Toda, N. Itoh, *J. Biosci. Bioeng.* **2012**, *113*, 12–19.
[84] S. Panke, M. G. Wubbolts, A. Schmid, B. Witholt, *Biotechnol. Bioeng.* **2000**, *69*, 91–100.
[85] H. Toda, R. Imae, N. Itoh, *Tetrahedron Asymmetry* **2012**, *23*, 1542–1549.
[86] E. Churakova, B. Tomaszewski, K. Buehler, A. Schmid, I. Arends, F. Hollmann, *Top. Catal.* **2014**, *57*, 385–391.
[87] F. Schulz, F. Leca, F. Hollmann, M. T. Reetz, *Beilstein J. Org. Chem.* **2005**, *1*, 10.
[88] H. Watanabe, H. Hirakawa, T. Nagamune, *ChemCatChem* **2013**, *5*, 3835–3840.
[89] A. Rioz-Martínez, G. de Gonzalo, D. E. T. Pazmiño, M. W. Fraaije, V. Gotor, *Eur. J. Org. Chem.* **2009**, *2009*, 2526–2532.
[90] C. E. Paul, I. W. Arends, F. Hollmann, *ACS Catal.* **2014**, *4*, 788–797.
[91] C. E. Paul, D. Tischler, A. Riedel, T. Heine, N. Itoh, F. Hollmann, *Acs Catal.* **2015**, *5*, 2961–2965.
[92] C. E. Paul, F. Hollmann, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 4773–4778.
[93] R.-E. Parker, N. S. Isaacs, *Chem. Rev.* **1959**, *59*, 737–799.
[94] A. Padwa, S. S. Murphree, *Arkivoc* **2006**, *3*, 6–33.
[95] C. Wang, L. Luo, H. Yamamoto, *Acc. Chem. Res.* **2016**, *49*, 193–204.
[96] W. A. Nugent, *J. Am. Chem. Soc.* **1992**, *114*, 2768–2769.
[97] G. J. Poelarends, J. E. van Hylckama Vlieg, J. R. Marchesi, L. M. Freitas Dos Santos, D. B. Janssen, *J. Bacteriol.* **1999**, *181*, 2050–2058.
[98] A. J. Van Den Wijngaard, D. B. Janssen, B. Witholt, *Microbiology* **1989**, *135*, 2199–2208.
[99] A. J. van den Wijngaard, P. T. Reuvekamp, D. B. Janssen, *J. Bacteriol.* **1991**, *173*, 124–129.
[100] J. H. Lutje Spelberg, L. Tang, M. van Gelder, R. M. Kellogg, D. B. Janssen, *Tetrahedron Asymmetry* **2002**, *13*, 1083–1089.
[101] E. J. de Vries, D. B. Janssen, *Curr. Opin. Biotechnol.* **2003**, *14*, 414–420.

[102] M. Schallmey, J. Koopmeiners, E. Wells, R. Wardenga, A. Schallmey, *Appl. Environ. Microbiol.* **2014**, *80*, 7303–7315.
[103] A. Schallmey, M. Schallmey, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 7827–7839.
[104] J. Koopmeiners, B. Halmschlag, M. Schallmey, A. Schallmey, *Appl. Microbiol. Biotechnol.* **2016**, 1–11.
[105] J. Koopmeiners, C. Diederich, J. Solarczek, H. Voß, J. Mayer, W. Blankenfeldt, A. Schallmey, *ACS Catal.* **2017**, *7*, 6877–6886.
[106] M. Schallmey, R. J. Floor, B. Hauer, M. Breuer, P. A. Jekel, H. J. Wijma, B. W. Dijkstra, D. B. Janssen, *ChemBioChem* **2013**, *14*, 870–881.
[107] C. Guo, Y. Chen, Y. Zheng, W. Zhang, Y. Tao, J. Feng, L. Tang, *Appl. Environ. Microbiol.* **2015**, *81*, 2919–2926.
[108] X. Wang, H. Lin, Y. Zheng, J. Feng, Z. Yang, L. Tang, *J. Biotechnol.* **2015**, *206*, 1–7.
[109] F. Watanabe, F. Yu, A. Ohtaki, Y. Yamanaka, K. Noguchi, M. Yohda, M. Odaka, *Proteins* **2015**, DOI 10.1002/prot.24938.
[110] D. B. Janssen, I. J. T. Dinkla, G. J. Poelarends, P. Terpstra, *Environ. Microbiol.* **2005**, *7*, 1868–1882.
[111] M. M. Elenkov, B. Hauer, D. B. Janssen, *Adv. Synth. Catal.* **2006**, *348*, 579–585.
[112] S.-W. Chen, S. S. Thakur, W. Li, C.-K. Shin, R. B. Kawthekar, G.-J. Kim, *J. Mol. Catal. Chem.* **2006**, *259*, 116–120.
[113] S. C. Davis, J. H. Grate, D. R. Gray, J. M. Gruber, G. W. Huisman, S. K. Ma, L. M. Newman, R. Sheldon, L. A. Wang, *Enzymatic Processes for the Production of 4-Substituted 3-Hydroxybutyric Acid Derivatives and Vicinal Cyano, Hydroxy Substituted Carboxylic Acid Esters*, **2005**, CA2535353 A1.
[114] S. K. Ma, J. Gruber, C. Davis, L. Newman, D. Gray, A. Wang, J. Grate, G. W. Huisman, R. A. Sheldon, *Green Chem.* **2010**, *12*, 81–86.
[115] N.-W. Wan, Z.-Q. Liu, F. Xue, Z.-Y. Shen, Y.-G. Zheng, *ChemCatChem* **2015**, *7*, 2446–2450.
[116] C. A. Jackevicius, M. M. Chou, J. S. Ross, N. D. Shah, H. M. Krumholz, *N. Engl. J. Med.* **2012**, *366*, 201–204.
[117] J. Halpern, *J. Phys. Chem.* **1959**, *63*, 398–403.
[118] M. Lamblin, L. Nassar-Hardy, J.-C. Hierso, E. Fouquet, F.-X. Felpin, *Adv. Synth. Catal.* **2010**, *352*, 33–79.
[119] M. D. Navalikhina, O. V. Krylov, *Russ. Chem. Rev.* **1998**, *67*, 587–616.
[120] Y. Monguchi, T. Ichikawa, H. Sajiki, *Chem. Pharm. Bull. (Tokyo)* **2017**, *65*, 2–9.
[121] S. C. Bergmeier, *Tetrahedron* **2000**, *56*, 2561–2576.
[122] T. Cedervall, I. Lynch, S. Lindman, T. Berggård, E. Thulin, H. Nilsson, K. A. Dawson, S. Linse, *Proc. Natl. Acad. Sci.* **2007**, *104*, 2050–2055.
[123] E. C. Theil, *Annu. Rev. Biochem.* **1987**, *56*, 289–315.
[124] S. Abe, J. Niemeyer, M. Abe, Y. Takezawa, T. Ueno, T. Hikage, G. Erker, Y. Watanabe, *J. Am. Chem. Soc.* **2008**, *130*, 10512–10514.
[125] S. Kanbak-Aksu, M. N. Hasan, W. R. Hagen, F. Hollmann, D. Sordi, R. A. Sheldon, I. Arends, *Chem. Commun.* **2012**, *48*, 5745–5747.
[126] A. S. Bommarius, J. K. Blum, M. J. Abrahamson, *Curr. Opin. Chem. Biol.* **2011**, *15*, 194–200.
[127] S. J. Kan, X. Huang, Y. Gumulya, K. Chen, F. H. Arnold, *Nature* **2017**, *552*, 132.
[128] S. J. Kan, R. D. Lewis, K. Chen, F. H. Arnold, *Science* **2016**, *354*, 1048–1051.
[129] S. K. Ma, J. Gruber, C. Davis, L. Newman, D. Gray, A. Wang, J. Grate, G. W. Huisman, R. A. Sheldon, *Green Chem.* **2010**, *12*, 81–86.
[130] J. C. Francis, P. E. Hansche, *Genetics* **1972**, *70*, 59–73.

[131] L. Pritchard, D. Corne, D. Kell, J. Rowland, M. Winson, *J. Theor. Biol.* **2005**, *234*, 497–509.
[132] R. Martínez, U. Schwaneberg, *Biol. Res.* **2013**, *46*, 395–405.
[133] S. Brakmann, *ChemBioChem* **2001**, *2*, 865–871.
[134] C. Dobson, W. Dall'Acqua, in *Methods Princ. Med. Chem.* (Eds.: T. Vaughan, J. Osbourn, B. Jallal), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, **2017**, pp. 189–227.
[135] U. T. Bornscheuer, M. Pohl, *Curr. Opin. Chem. Biol.* **2001**, *5*, 137–143.
[136] M. T. Reetz, D. Kahakeaw, R. Lohmer, *ChemBioChem* **2008**, *9*, 1797–1804.
[137] P. Larrañaga, B. Calvo, R. Santana, C. Bielza, J. Galdiano, I. Inza, J. A. Lozano, R. Armañanzas, G. Santafé, A. Pérez, et al., *Brief. Bioinform.* **2006**, *7*, 86–112.
[138] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, *Nat. Protoc.* **2015**, *10*, 845–858.
[139] P. Jones, D. Binns, H.-Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, et al., *Bioinformatics* **2014**, *30*, 1236–1240.
[140] M. Musil, J. Stourac, J. Bendl, J. Brezovsky, Z. Prokop, J. Zendulka, T. Martinek, D. Bednar, J. Damborsky, *Nucleic Acids Res.* **2017**, *45*, W393–W399.
[141] R. K. Kuipers, H.-J. Joosten, W. J. van Berkel, N. G. Leferink, E. Rooijen, E. Ittmann, F. van Zimmeren, H. Jochens, U. Bornscheuer, G. Vriend, *Proteins Struct. Funct. Bioinforma.* **2010**, *78*, 2101–2113.
[142] R. Chaloupkova, L. Sumbalova, J. Stourac, T. Martinek, D. Bednar, **2017**.
[143] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639–1662.
[144] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455–461.
[145] Schrödinger, LLC, **2015**.
[146] K. W. Kaufmann, G. H. Lemmon, S. L. DeLuca, J. H. Sheehan, J. Meiler, *Biochemistry* **2010**, *49*, 2987–2998.
[147] I. W. Davis, W. B. Arendall III, D. C. Richardson, J. S. Richardson, *Structure* **2006**, *14*, 265–274.
[148] N. Metropolis, S. Ulam, *J. Am. Stat. Assoc.* **1949**, *44*, 335–341.
[149] N. Ollikainen, R. M. de Jong, T. Kortemme, *PLOS Comput Biol* **2015**, *11*, e1004335.
[150] J. A. G. M. de Visser, T. F. Cooper, S. F. Elena, *Proc. R. Soc. B Biol. Sci.* **2011**, *278*, 3617–3624.
[151] C. M. Miton, N. Tokuriki, *Protein Sci.* **2016**, *25*, 1260–1272.
[152] R. Sharma, P. G. Bulger, M. McNevin, P. G. Dormer, R. G. Ball, E. Streckfuss, J. F. Cuff, J. Yin, C. Chen, *Org. Lett.* **2009**, *11*, 3194–3197.
[153] S. E. Denmark, Z. Wu, C. M. Crudden, H. Matsuhashi, *J. Org. Chem.* **1997**, *62*, 8288–8289.
[154] W. Kroutil, M. Mischitz, K. Faber, *J. Chem. Soc. Perkin 1* **1997**, *0*, 3629–3636.
[155] E. G. Ankudey, H. F. Olivo, T. L. Peeples, *Green Chem.* **2006**, *8*, 923–926.
[156] J. Lindborg, A. Tanskanen, L. T. Kanerva, *Biocatal. Biotransformation* **2009**, *27*, 204–210.
[157] G. Righi, S. Pietrantonio, C. Bonini, *Tetrahedron* **2001**, *57*, 10039–10046.
[158] P. M. G. Bavin, *Org. Synth.* **1960**, *40*, 5.
[159] A. Schallmey, G. den Besten, I. G. P. Teune, R. F. Kembaren, D. B. Janssen, *Appl. Microbiol. Biotechnol.* **2010**, *89*, 1475–1485.
[160] P. Bracco, D. B. Janssen, A. Schallmey, *Microb. Cell Factories* **2013**, *12*, 95.
[161] L. Wiermans, S. Hofzumahaus, C. Schotten, L. Weigand, M. Schallmey, A. Schallmey, P. Domínguez de María, *ChemCatChem* **2013**, *5*, 3719–3724.

[162] R. M. Haak, C. Tarabiono, D. B. Janssen, A. J. Minnaard, J. G. de Vries, B. L. Feringa, *Org. Biomol. Chem.* **2007**, *5*, 318–323.
[163] D. Hanahan, *J. Mol. Biol.* **1983**, *166*, 557–580.
[164] J. A. Meyers, D. Sanchez, L. P. Elwell, S. Falkow, *J. Bacteriol.* **1976**, *127*, 1529–1537.
[165] R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, H. A. Erlich, *Science* **1988**, *239*, 487–491.
[166] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, H. Erlich, *Cold Spring Harb. Symp. Quant. Biol.* **1986**, *51*, 263–273.
[167] J. Koopmeiners, B. Halmschlag, M. Schallmey, A. Schallmey, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 7517–7527.
[168] H. Schägger, G. von Jagow, *Anal. Biochem.* **1987**, *166*, 368–379.
[169] T. Omura, R. Sato, *J Biol Chem* **1964**, *239*, 2370–2378.
[170] T. Lacour, T. Achstetter, B. Dumas, *J. Biol. Chem.* **1998**, *273*, 23984–23992.
[171] K. A. Johnson, R. S. Goody, *Biochemistry* **2011**, *50*, 8264–8269.
[172] A. V. Hill, *J Physiol* **1910**, *40*, 4–7.
[173] M. I. Stefan, N. L. Novere, *PLoS Comput. Biol.* **2013**, *9*, 1003106.
[174] C. S. Chen, Y. Fujimoto, G. Girdaukas, C. J. Sih, *J. Am. Chem. Soc.* **1982**, *104*, 7294–7299.
[175] B. Yang, Z. Lu, *ACS Catal.* **2017**, *7*, 8362–8365.
[176] I. A. Sayyed, A. Sudalai, *Tetrahedron Asymmetry* **2004**, *15*, 3111–3116.
[177] Z. Findrik, A. V. Presečki, Đ. Vasić-Rački, *Biochem. Eng. J.* **2012**, *60*, 91–98.
[178] M. A. Miteva, F. Guyon, P. Tufféry, *Nucleic Acids Res.* **2010**, *38*, 622–627.
[179] G. Crooks, G. Hon, J. Chandonia, S. Brenner, *Genome Res* **2004**, *14*, 1188–1190.
[180] F. W. Studier, B. A. Moffatt, *J. Mol. Biol.* **1986**, *189*, 113–130.
[181] K. Herzog, P. Bracco, A. Onoda, T. Hayashi, K. Hoffmann, A. Schallmey, *Biol. Crystallogr.* **2014**, *70*.
[182] G. Hasnaoui-Dijoux, M. Majerić Elenkov, J. H. Lutje Spelberg, B. Hauer, D. B. Janssen, *ChemBioChem* **2008**, *9*, 1048–1051.
[183] E. Calderini, J. Wessel, P. Süss, P. Schrepfer, R. Wardenga, A. Schallmey, *ChemCatChem* **2019**, *11*, 2099–2106.
[184] R. M. de Jong, J. J. W. Tiesinga, Rozeboom, H.J., Kalk, K.H., Tang, L., Janssen, D.B., Dijkstra, B.W., *EMBO J.* **2003**, *22*, 4933–4944.
[185] R. M. de Jong, J. J. W. Tiesinga, A. Villa, L. Tang, D. B. Janssen, B. W. Dijkstra, *J. Am. Chem. Soc.* **2005**, *127*, 13338–13343.
[186] **N.d.**
[187] Y. Shi, *Acc. Chem. Res.* **2004**, *37*, 488–496.
[188] R. C. Thompson, P. Wieland, E. H. Appelman, *Inorg. Chem.* **1979**, *18*, 1974–1977.
[189] K.-Y. A. Lin, Y.-C. Chen, *J. Taiwan Inst. Chem. Eng.* **2016**, *60*, 423–429.
[190] D. Méndez-Sánchez, N. Ríos-Lombardía, V. Gotor, V. Gotor-Fernández, *Tetrahedron* **2014**, *70*, 1144–1148.
[191] M. Kafri, E. Metzl-Raz, G. Jona, N. Barkai, *Cell Rep.* **2016**, *14*, 22–31.
[192] G. L. Rosano, E. A. Ceccarelli, *Front. Microbiol.* **2014**, *5*, DOI 10.3389/fmicb.2014.00172.
[193] S. Garneau-Tsodikova, K. J. Labby, *MedChemComm* **2016**, *7*, 11–27.
[194] F.-A. Delmani, A. S. Jaran, Y. A. Tarazi, H. Masaadeh, O. Zaki, *Asian J. Biomed. Pharm. Sci.* **2017**, *7*.
[195] M.-P. Mingeot-Leclercq, Y. Glupczynski, P. M. Tulkens, *Antimicrob. Agents Chemother.* **1999**, *43*, 727–737.

[196] M. Kluge, R. Ullrich, K. Scheibner, M. Hofrichter, *Green Chem.* **2012**, *14*, 440–446.
[197] A. Kumar, K. Dhar, S. S. Kanwar, P. K. Arora, *Biol. Proced. Online* **2016**, *18*, 2.
[198] J. S. Dordick, *Biotechnol. Prog.* **1992**, *8*, 259–267.
[199] H. Toda, T. Ohuchi, R. Imae, N. Itoh, *Appl. Environ. Microbiol.* **2015**, *81*, 1919–1925.
[200] J. Volmer, A. Schmid, B. Bühler, *Biotechnol. J.* **2017**, *12*, 1600558.
[201] M. L. Corrado, T. Knaus, F. G. Mutti, *ChemBioChem* **2018**, *19*, 679–686.
[202] K. M. Bekers, J. J. Heijnen, W. M. van Gulik, *Yeast* **2015**, *32*, 541–557.
[203] L. Shu, Y. Shi, *Tetrahedron* **2001**, *57*, 5213–5218.
[204] M. An, W. Liu, X. Zhou, R. Ma, H. Wang, B. Cui, W. Han, N. Wan, Y. Chen, *RSC Adv.* **2019**, *9*, 16418–16422.
[205] R. M. de Jong, K. H. Kalk, L. Tang, D. B. Janssen, B. W. Dijkstra, *J. Bacteriol.* **2006**, *188*, 4051–4056.
[206] F. Watanabe, F. Yu, A. Ohtaki, Y. Yamanaka, K. Noguchi, M. Yohda, M. Odaka, *Proteins Struct. Funct. Bioinforma.* **2015**, *83*, 2230–2239.
[207] E. J. Land, W. A. Prütz, *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.* **1979**, *36*, 75–83.
[208] Q. Liao, X. Du, W. Jiang, Y. Tong, Z. Zhao, R. Fang, J. Feng, L. Tang, *J. Biotechnol.* **2018**, *272–273*, 48–55.
[209] S.-P. Zou, Y.-G. Zheng, E.-H. Du, Z.-C. Hu, *J. Biotechnol.* **2014**, *188*, 42–47.
[210] C. P. Adams, K. A. Walker, S. O. Obare, K. M. Docherty, *PLOS ONE* **2014**, *9*, e85981.
[211] J. H. Schrittwieser, F. Coccia, S. Kara, B. Grischek, W. Kroutil, N. d'Alessandro, F. Hollmann, *Green Chem.* **2013**, *15*, 3318.
[212] H. Arabnejad, M. Dal Lago, P. A. Jekel, R. J. Floor, A.-M. W. H. Thunnissen, A. C. Terwisscha van Scheltinga, H. J. Wijma, D. B. Janssen, *Protein Eng. Des. Sel.* **2017**, *30*, 175–189.
[213] H. Sato, W. Hummel, H. Gröger, *Angew. Chem. Int. Ed.* **2015**, *54*, 4488–4492.
[214] G. Sabitha, R. S. Babu, M. S. K. Reddy, J. S. Yadav, *Synthesis* **2002**, *2002*, 2254–2258.
[215] A. Freeman, J. M. Woodley, M. D. Lilly, *Bio/Technology* **1993**, *11*, 1007.
[216] M. A. Moreira, T. B. Bitencourt, M. da G. Nascimento, *Synth. Commun.* **2005**, *35*, 2107–2114.
[217] P. J. Ziemann, *J. Phys. Chem. A* **2002**, *106*, 4390–4402.
[218] J. Lee, N. M. Goodey, *Chem. Rev.* **2011**, *111*, 7595–7624.
[219] K. E. Omari, S. Liekens, L. E. Bird, J. Balzarini, D. K. Stammers, *Mol. Pharmacol.* **2006**, *69*, 1891–1896.
[220] A. Aharoni, L. Gaidukov, O. Khersonsky, S. M. Gould, C. Roodveldt, D. S. Tawfik, *Nat. Genet.* **2005**, *37*, 73–76.

Appendix

List of tables

List of figures

List of abbreviations

AFHS	Apoferritin from horse spleen
Amp	ampicillin
APS	Ammonium persulfate
AIXC	Anion exchange chromatography
BNAH	1-benzyl-1,4-dihydronicotinamide
bp	Base pairs
BVMO	Baeyer-Villiger monooxygenase
CFE	Cell free extract
Cm	chloramphenicol
CV	Column volume
CYP	Cytochrome P450 monooxygenase
ddH_2O	Double distilled water
dO_2	dissolved oxygen

DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
e.g.	*exempli gratia* (for example)
et al.	*et alii (and others)*
EtOAc	Ethyl acetate
EtOH	Ethanol
EV	Empty vector, negative control
F	Flow-through
FID	Flame ionization detector
FPLC	Fast protein liquid chromatography
GC	Gas chromatography
GC-MS	Gas chromatography- mass spectroscopy
HHDH	Halohydrin dehalogenase
HIC	Hydrophobic interaction chromatography
HR-MS	High resolution - mass spectroscopy
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ISTD	Internal standard
kan	Kanamycin
kb	Kilobase
k_{cat}	Turnover number
kDa	Kilodalton
LB	Luria broth

M	Marker
MsAct	Acyl transferase from *Mycobacterium smegmatis*
Nu^-	Nucleophile
OD_{600}	Optical density at λ = 600 nm
PCR	Polymerase chain reaction
Pd/C	Palladium on carbon
PdR	Putidaredoxin reductase from *Pseudomonas putida*
Pdx	Putidaredoxin from *Pseudomonas putida*
PMS	Phenylmethylsulfonyl fluoride
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SOC	Super optimal broth with Catabolite repression
Spec	Spectinomycin
SSM	Site-saturation mutagenesis
StyAB	Styrene monooxygenase
TAE	Tris base, acetic acid and EDTA
TB	Terrific broth
TBME	Methyl *tert*-butyl ether
TE	Tris base, EDTA
TEM	Transmission electron microscopy
TEMED	Tetramethylethylenediamine
TES	Triethylsilane
TFAA	Trifluoroacetic anhydride
TLC	Thin Layer Chromatography
T_m	Melting temperature

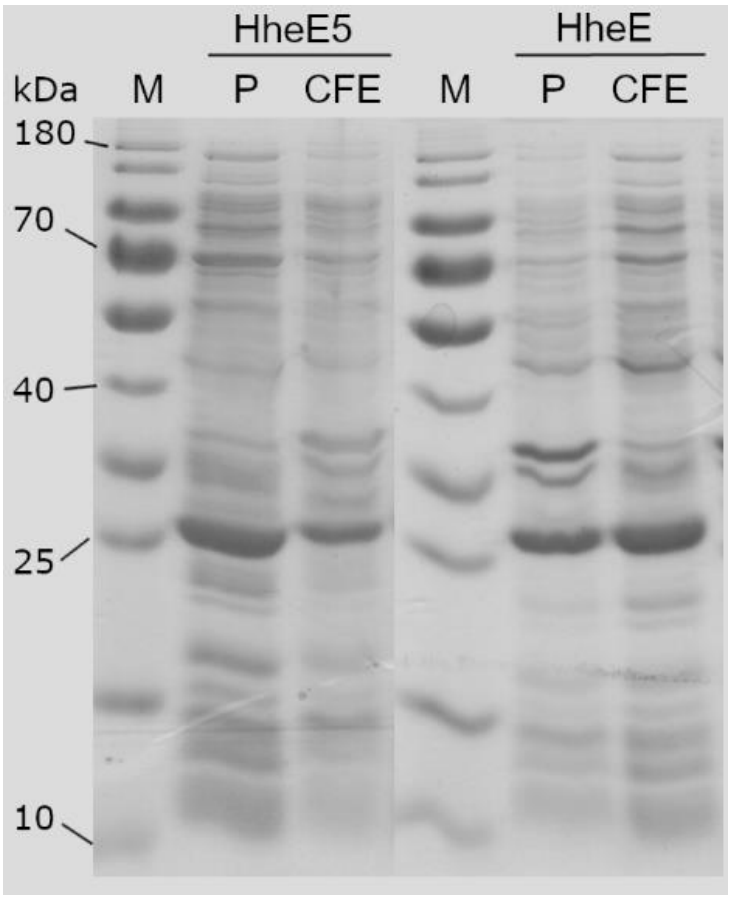
Tris	Tris(hydroxymethyl)aminomethane
UPO	Unspecific peroxygenase
UV	Ultraviolet
W	Wash fraction
WT	Wild type
X^-	Halide

Abbreviations of amino acids:

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Supplementary material

Heterologous expression



Supporting figure 1. Exemplary SDS-PAGE corresponding to fermentations for the production of HheE5 or HheE. P corresponds to the insoluble fraction and CFE corresponds to the soluble fraction.

Epoxide ring opening supporting tables

Supporting table 1. Regiopreferences of active HHDHs in the azidolysis of (*R*,*R*)- and (*S*,*S*)-enantiomers of epoxide substrates **1**, **2** and **5**. Regiopreference is given as the ratio of nucleophilic attack at the □-carbon (sterically less hindered) and the βcarbon (sterically more hindered).

	Regiopreference (α:β attack)					
HHDH	(R,R)-**1**	(S,S)-**1**	(R,R)-**2**	(S,S)-**2**	(R,R)-**5**	(S,S)-**5**
HheA3	12:1	12:1	21:1	13:1	n.d.[a]	n.d.[a]
HheB3	11:1	1:1	11:1	1:1	n.d.[a]	n.d.[a]
HheB5	67:1	3:1	162:1	2:1	n.d.[a]	n.d.[a]
HheB6	105:1	4:1	151:1	4:1	n.d.[a]	n.d.[a]
HheB7	153:1	4:1	161:1	3:1	n.d.[a]	n.d.[a]
HheC	60:1	18:1	332:1	26:1	n.d.[a]	n.d.[a]
HheD	8:1	1:3	16:1	1:4	n.d.[a]	n.d.[a]
HheD2	6:1	1:3	7:1	1:4	n.d.[a]	n.d.[a]
HheD3	15:1	1:5	14:1	1:6	1:7	1:80
HheD5	16:1	1:2	1:1	1:1	n.d.[a]	n.d.[a]
HheE	97:1	40:1	217:1	36:1	n.d.[a]	n.d.[a]
HheE2	9:1	5:1	55:1	20:1	n.d.[a]	n.d.[a]
HheE3	8:1	8:1	62:1	27:1	n.d.[a]	n.d.[a]
HheE4	n.d.[a]	n.d.[a]	6:1	11:1	n.d.[a]	n.d.[a]
HheE5	83:1	30:1	327:1	60:1	n.d.[a]	n.d.[a]
HheF	4:1	7:1	5:1	7:1	1:4	1:113
HheG	1:1	1:1	21:1	1:4	1:118	1:1743
HheG2	1:1	1:1	1:1	1:1	1:111	1:5702

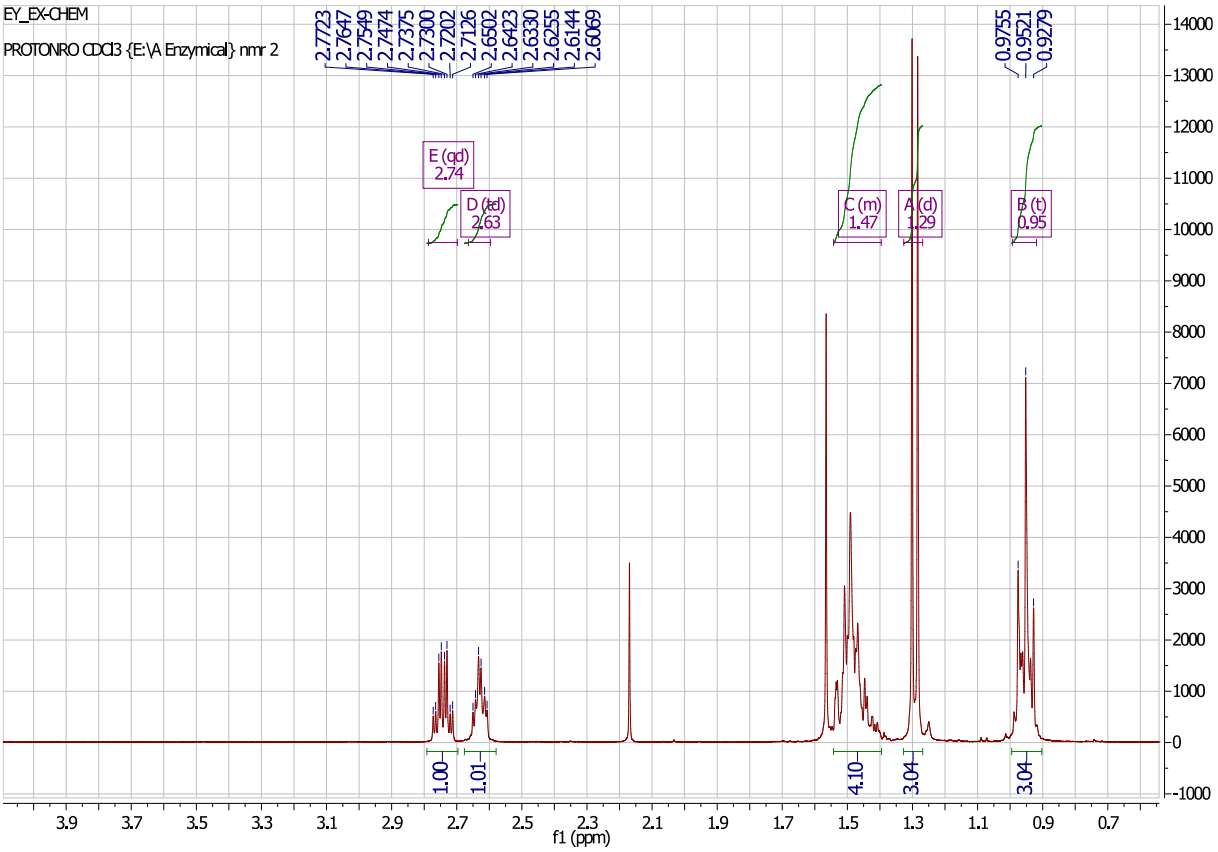
[a] not determined

Supporting table 2. Conversions (C), enantiomeric excesses (ee_P) and calculated apparent enantiomeric ratios (E_{app}) of regiomeric azidoalcohols **10a** and **b** formed in the HHDH-catalysed azidolysis of epoxide **5**. The absolute configuration of the preferentially formed enantiomer of each regioisomer is given in parentheses.

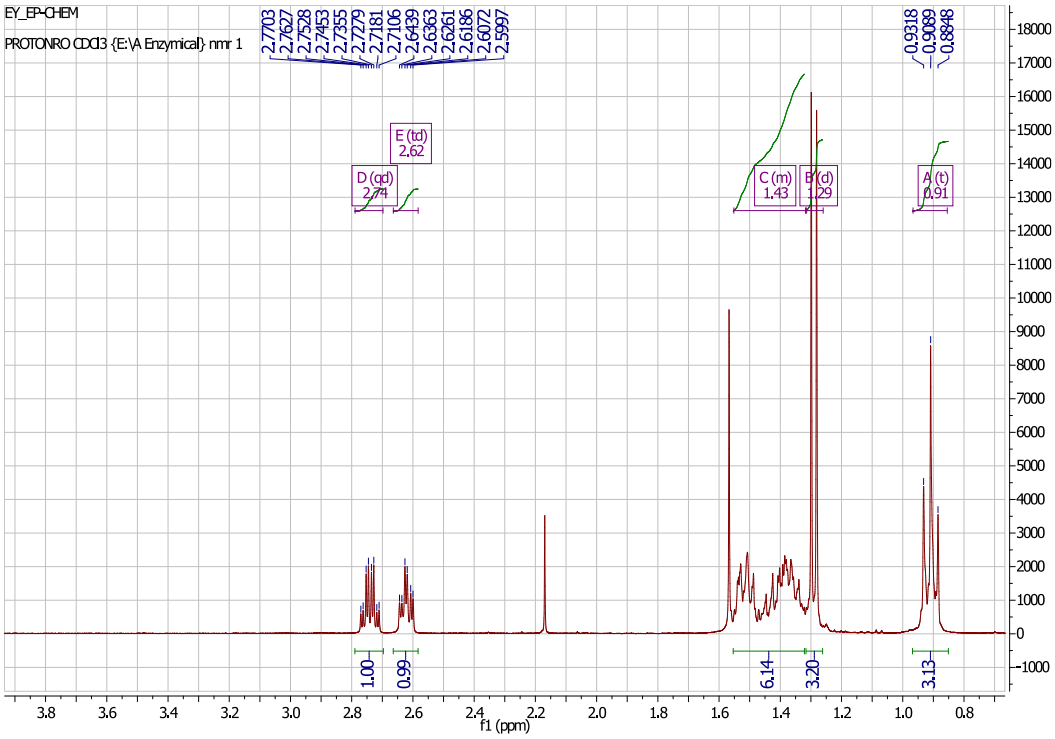
	10**a**			10**b**		
HHDH	C [%]	ee_P [%]	E^{app}	C [%]	ee_P [%]	E^{app}
HheD3	0.08	80	9.0 (R,S)	1.1	7.5	1.2 (R,S)
HheF	0.17	95	39 (S,R)	1.2	17	1.4 (S,R)
HheG	0.02	83	11 (R,S)	44	11	1.3 (R,S)
HheG2	0.01	93	28 (R,S)	24	26	1.8 (R,S)

^{1}H, DEPT, COSY and HSQC NMR spectra

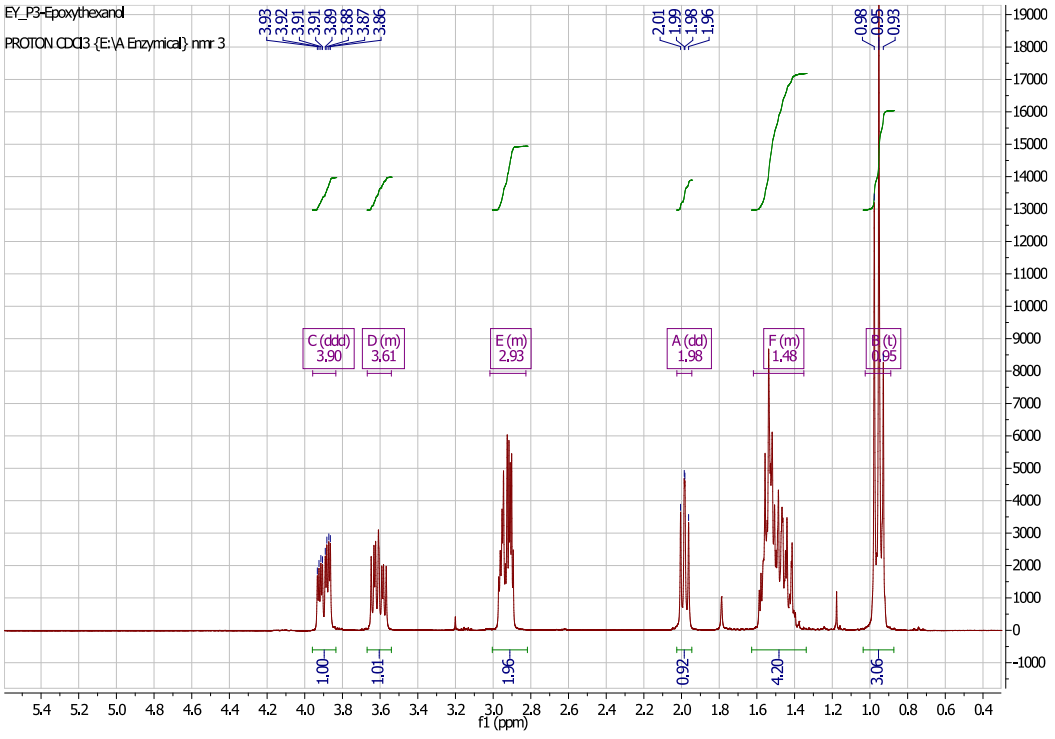
Epoxides



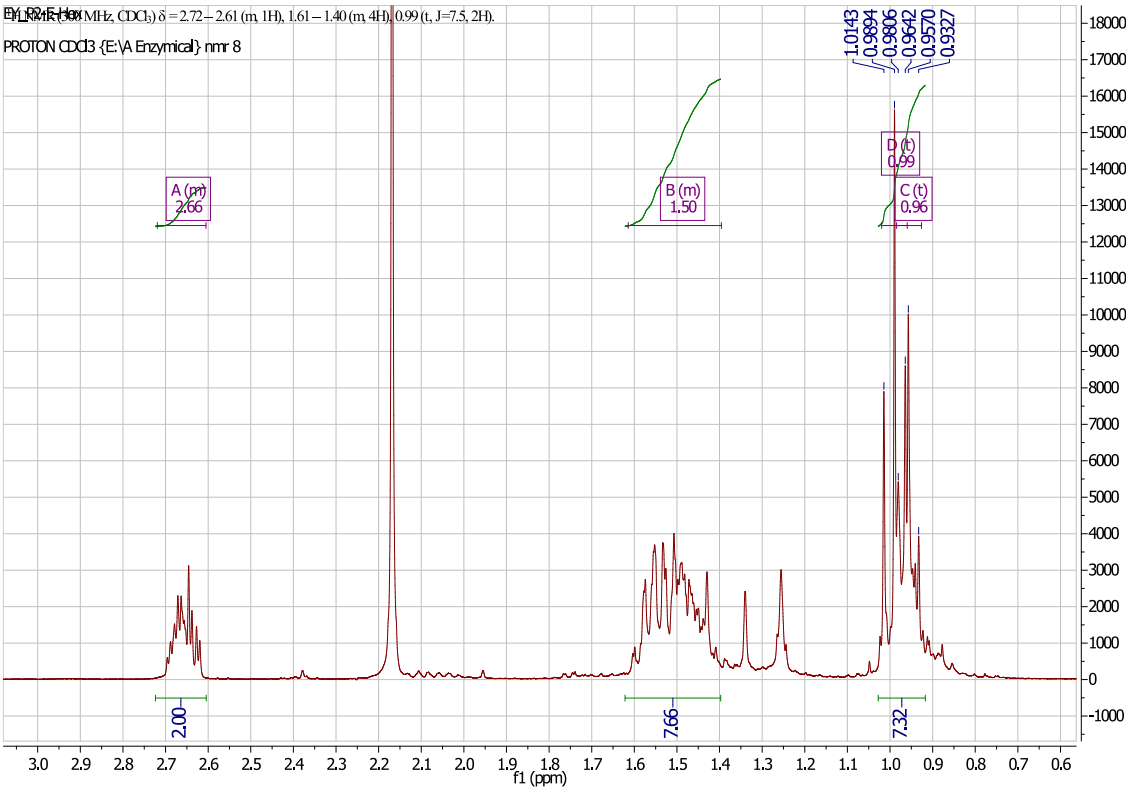
Supporting figure 2. ^{1}H NMR (300 MHz, $CDCl_3$) of *trans*-2,3-epoxyhexane (**21**) (Spectrum analysed with MestReNova 6).



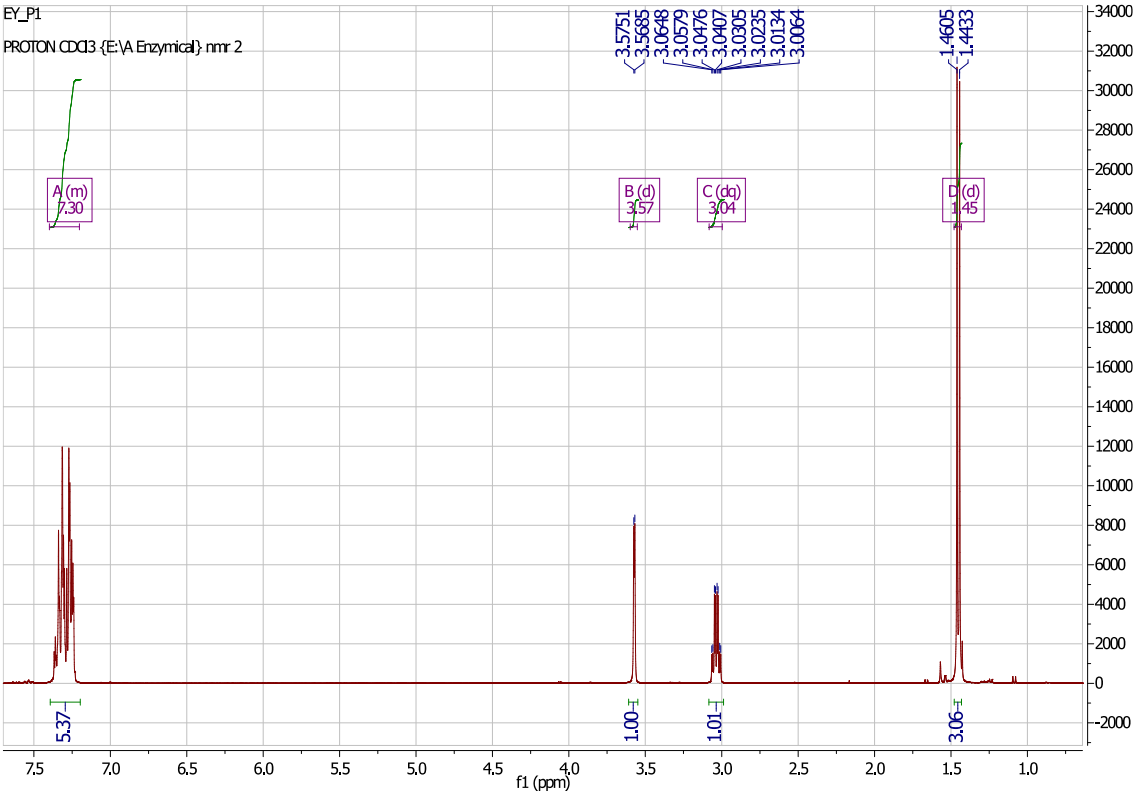
Supporting figure 3. ^{1}H NMR (300 MHz, $CDCl_3$) of *trans*-2,3-epoxyheptane (**22**) (Spectrum analysed with MestReNova 6).



Supporting figure 4. ^{1}H NMR (300 MHz, $CDCl_3$) of *trans*-4,5-epoxyhexanol (**23**) (Spectrum analysed with MestReNova 6).

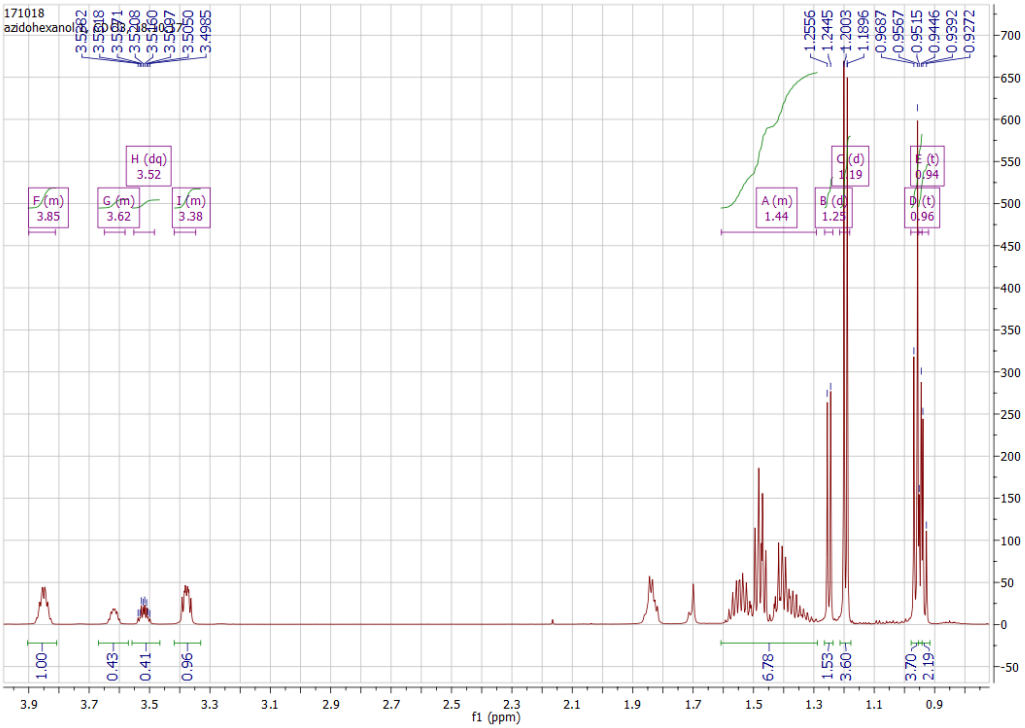


Supporting figure 5. ^{1}H NMR (300 MHz, $CDCl_3$) of *trans*-3,4-epoxyheptane (**24**) (Spectrum analysed with MestReNova 6).

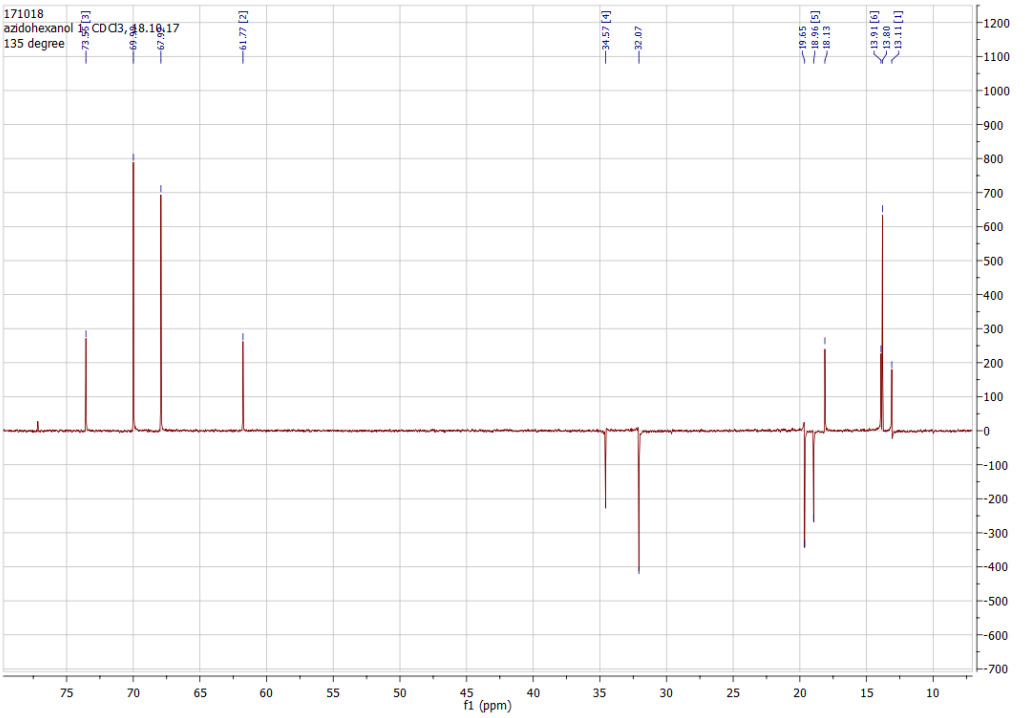


Supporting figure 6^{1}H NMR (300 MHz, $CDCl_3$) of *trans*-1-phenylpropylene oxide (**25**) (Spectrum analysed with MestReNova 6).

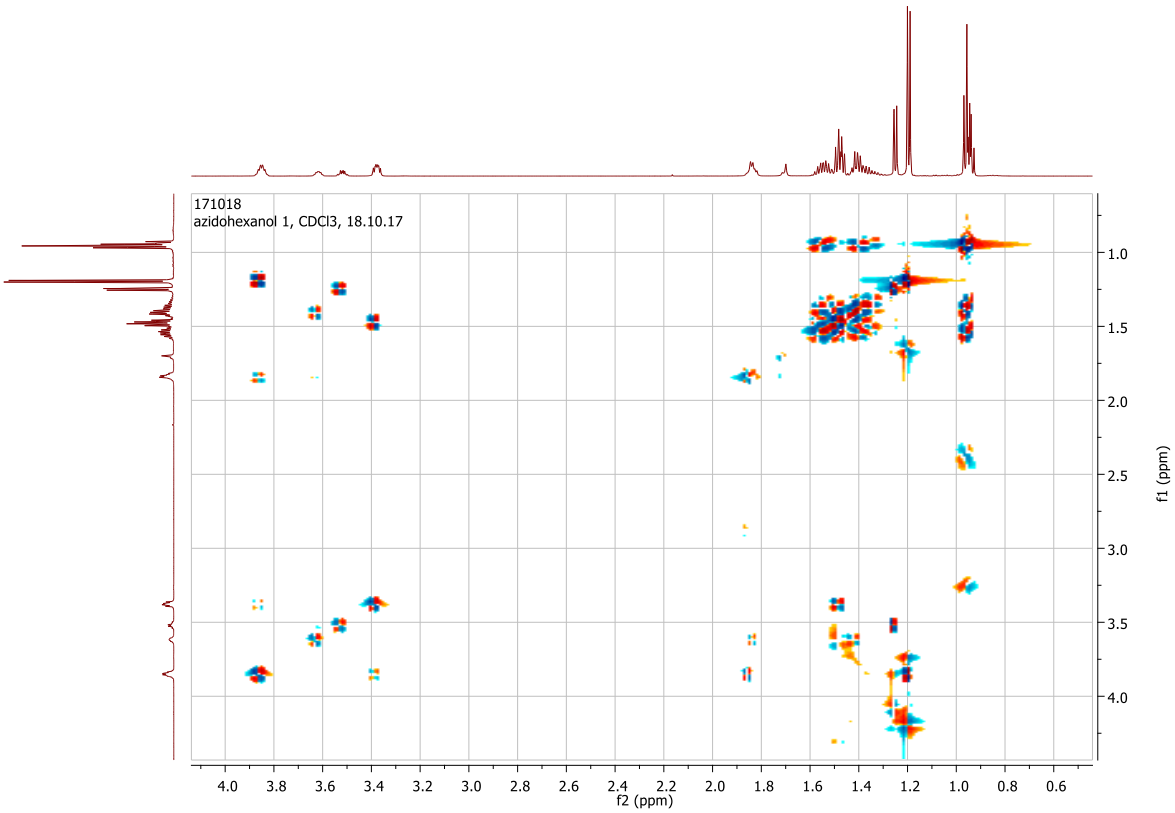
Azido alcohols



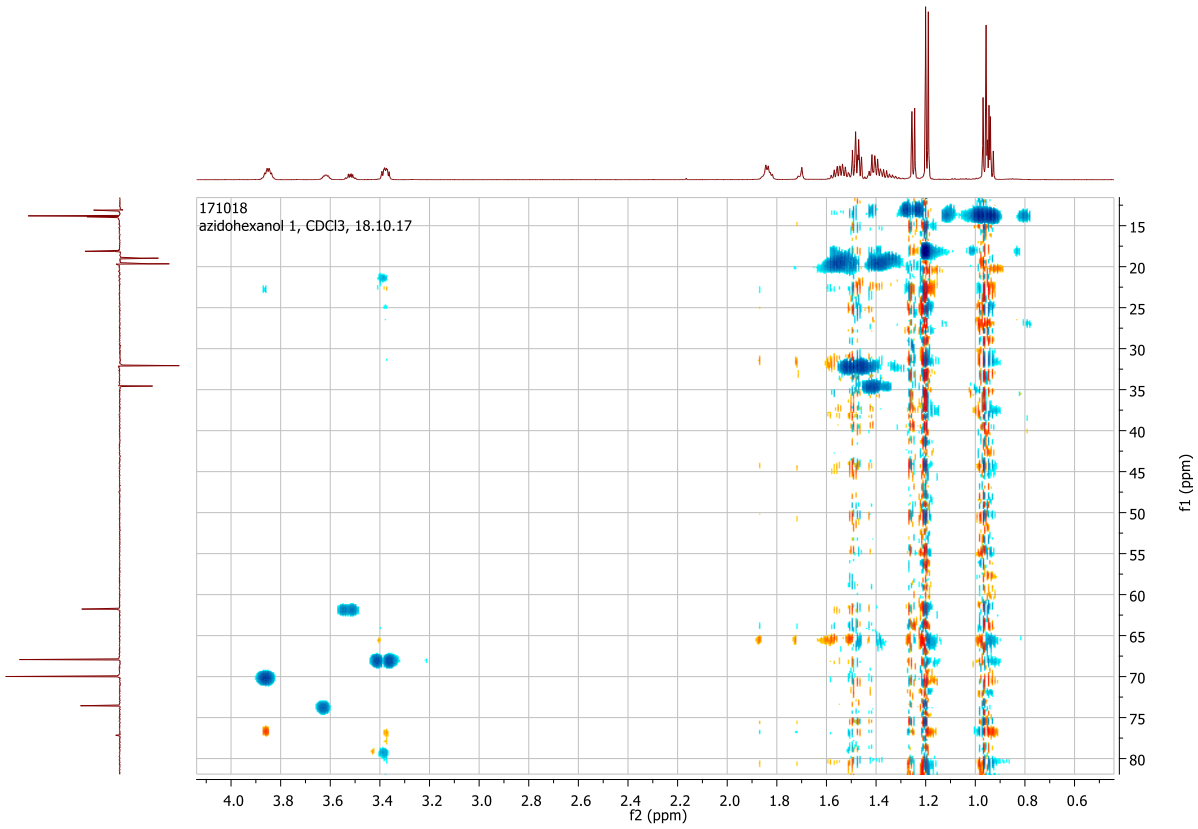
Supporting figure 7. ^{1}H NMR (600 MHz, $CDCl_3$) of the regioisomeric mixture of 2-azido-3-hydroxyhexane (**31a**) and 3-azido-2-hydroxyhexane (**31b**) (Spectrum analysed with MestReNova 6).



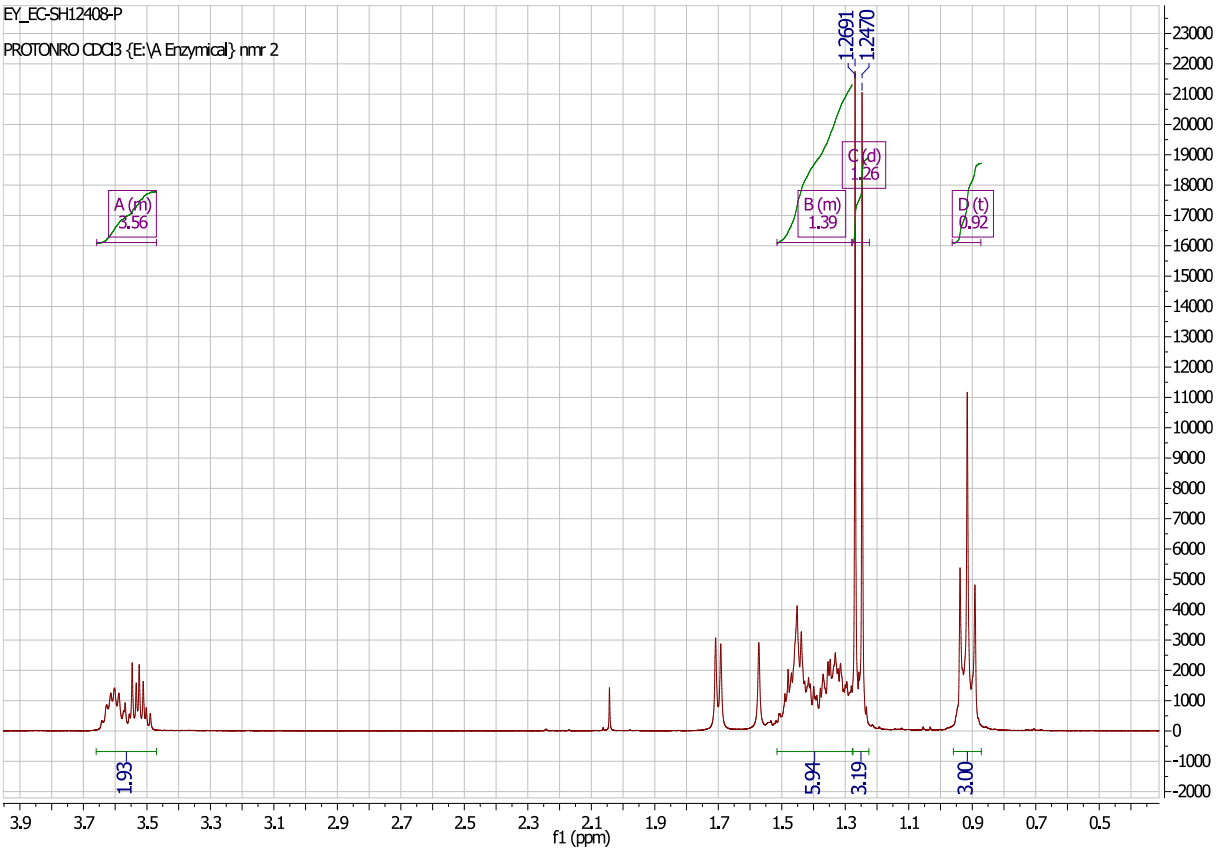
Supporting figure 8. DEPT (151 MHz, $CDCl_3$) of the regioisomeric mixture of 2-azido-3-hydroxyhexane (**31a**) and 3-azido-2-hydroxyhexane (**31b**) (Spectrum analysed with MestReNova 6).



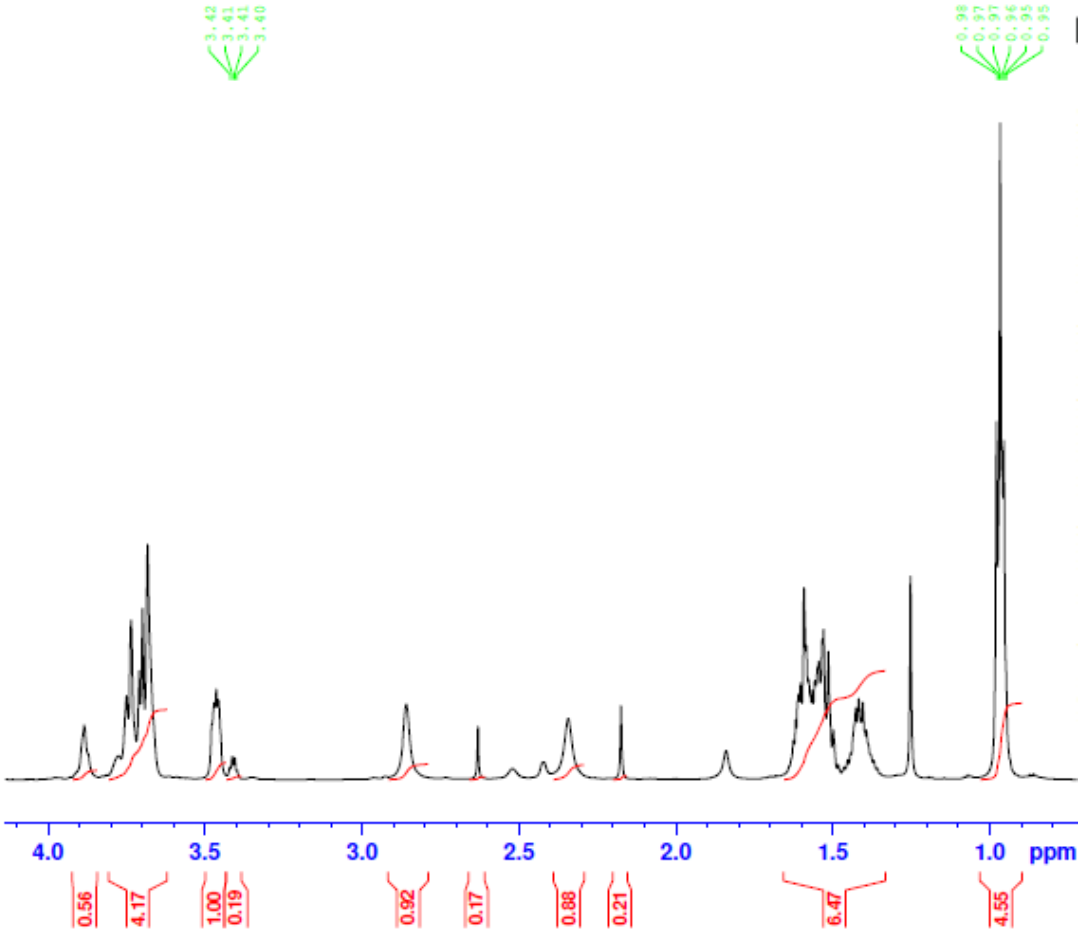
Supporting figure 9. COSY spectrum with double quantum filter, applied for the identification of regioisomers of 2-azido-3-hydroxyhexane (**31a**) and 3-azido-2-hydroxyhexane (**31b**) (Spectrum analysed with MestReNova 6).



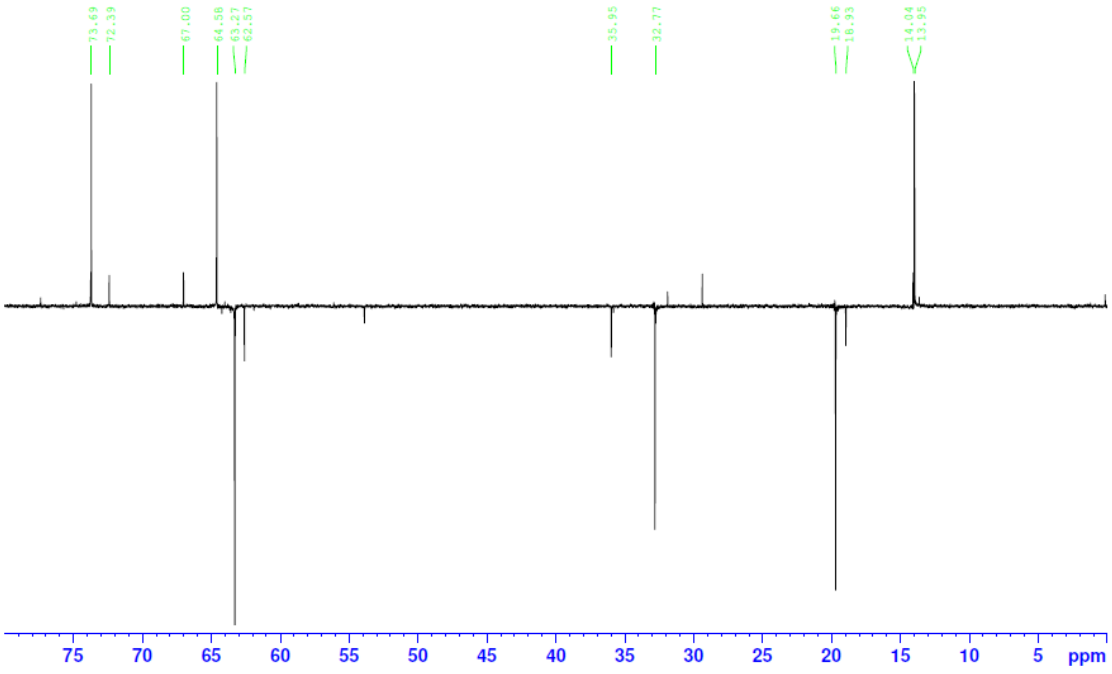
Supporting figure 10. HSQC spectrum, applied for the identification of regioisomers of 2-azido-3-hydroxyhexane (**31a**) and 3-azido-2-hydroxyhexane (**31b**) (Spectrum analysed with MestReNova 6).



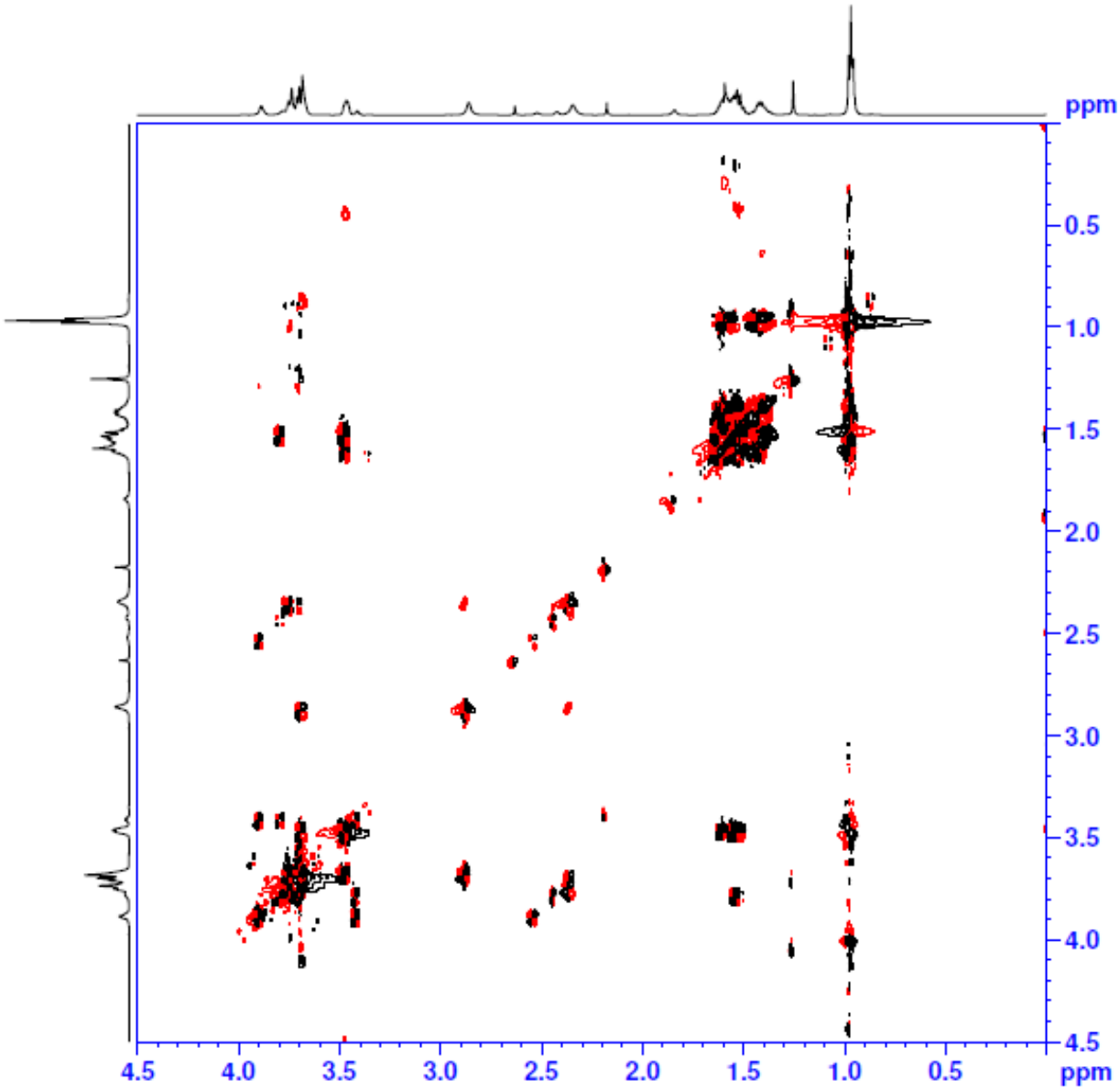
Supporting figure 11. ^{1}H NMR (600 MHz, $CDCl_3$) of the regioisomeric mixture of 2-azido-3-hydroxyheptane (**32a**) and 3-azido-2-hydroxyheptane (**32b**) (Spectrum analysed with MestReNova 6).



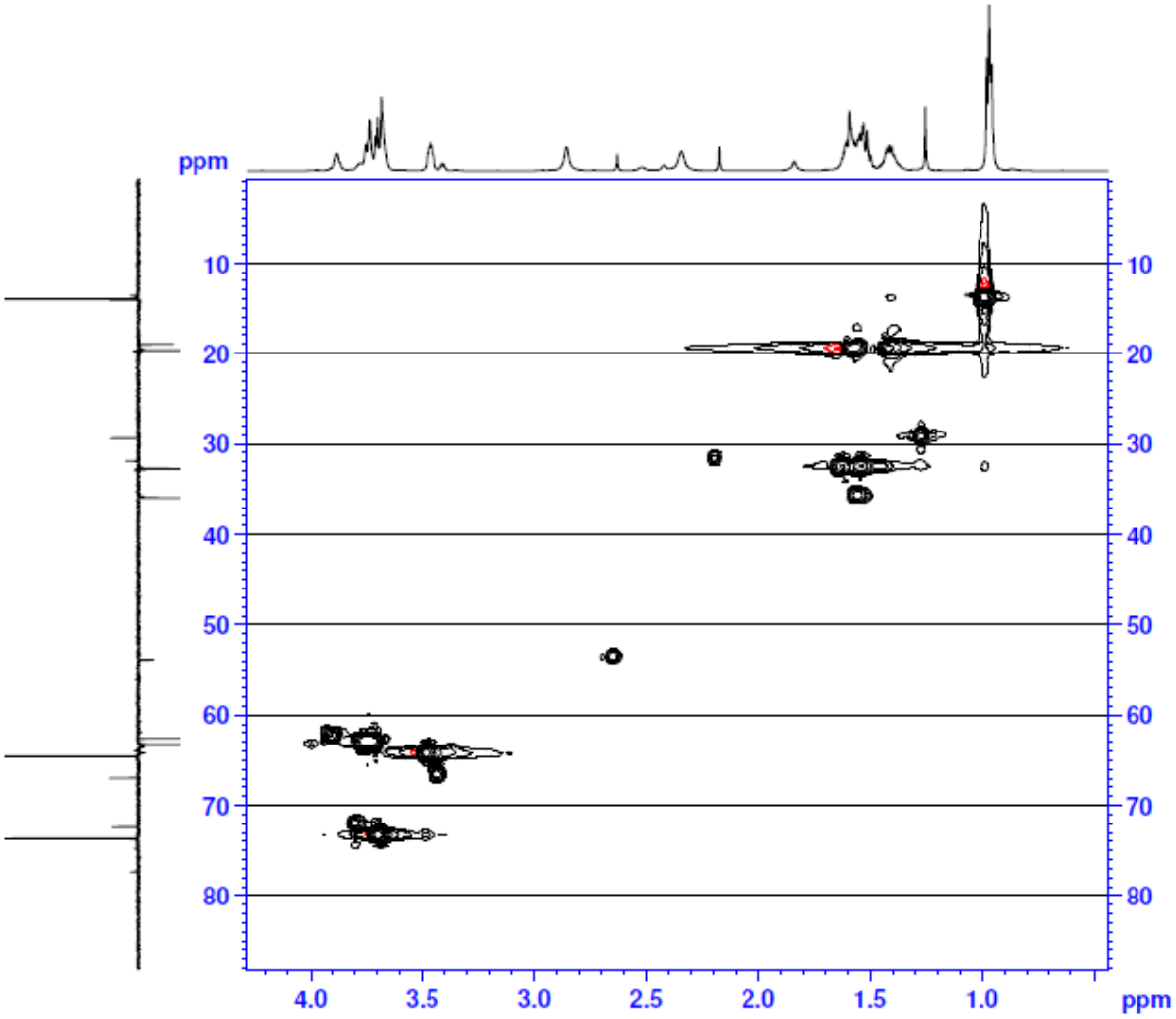
Supporting figure 12. ^{1}H NMR (600 MHz, $CDCl_3$) of the regioisomeric mixture of azidohexan-diol **33** (Spectrum analysed with TopSpin 4.0).



Supporting figure 13. DEPT (151 MHz, $CDCl_3$) of regioisomeric mixture of azidohexan-diol **33** (Spectrum analysed with TopSpin 4.0).

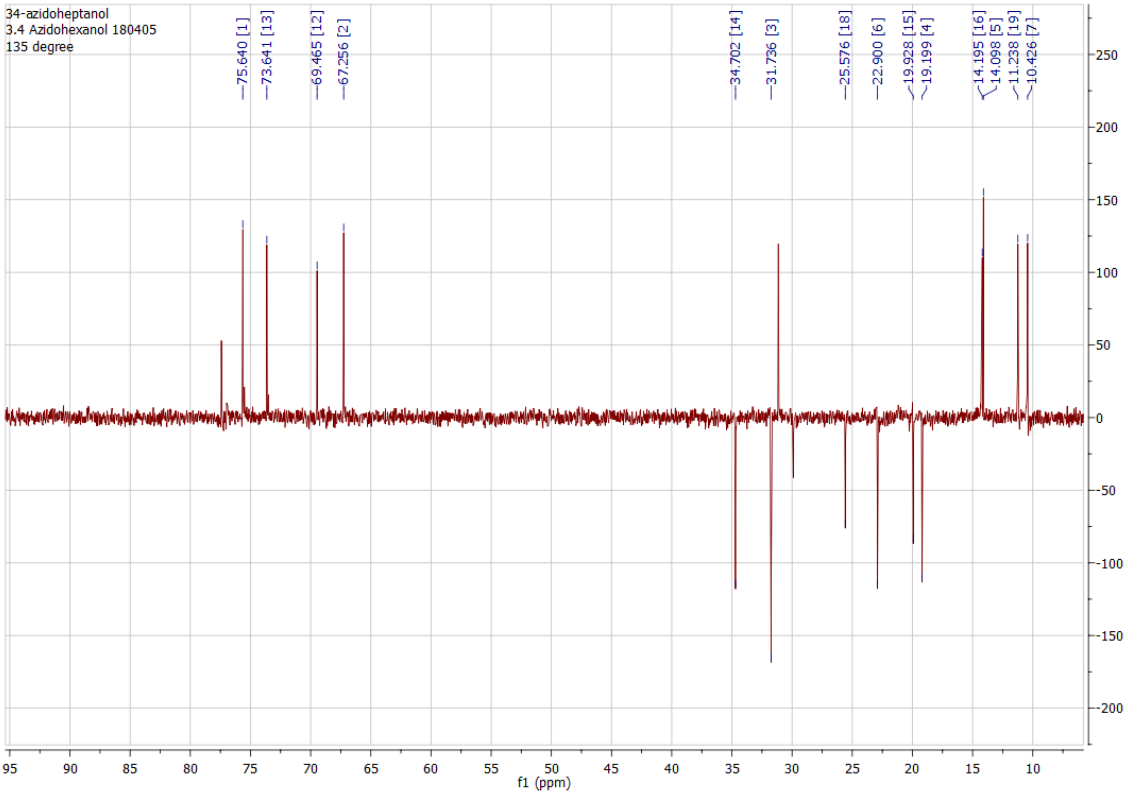


Supporting figure 14. COSY spectrum with double quantum filter, applied for the identification of regioisomers of **33** (Spectrum analysed with TopSpin 4.0).

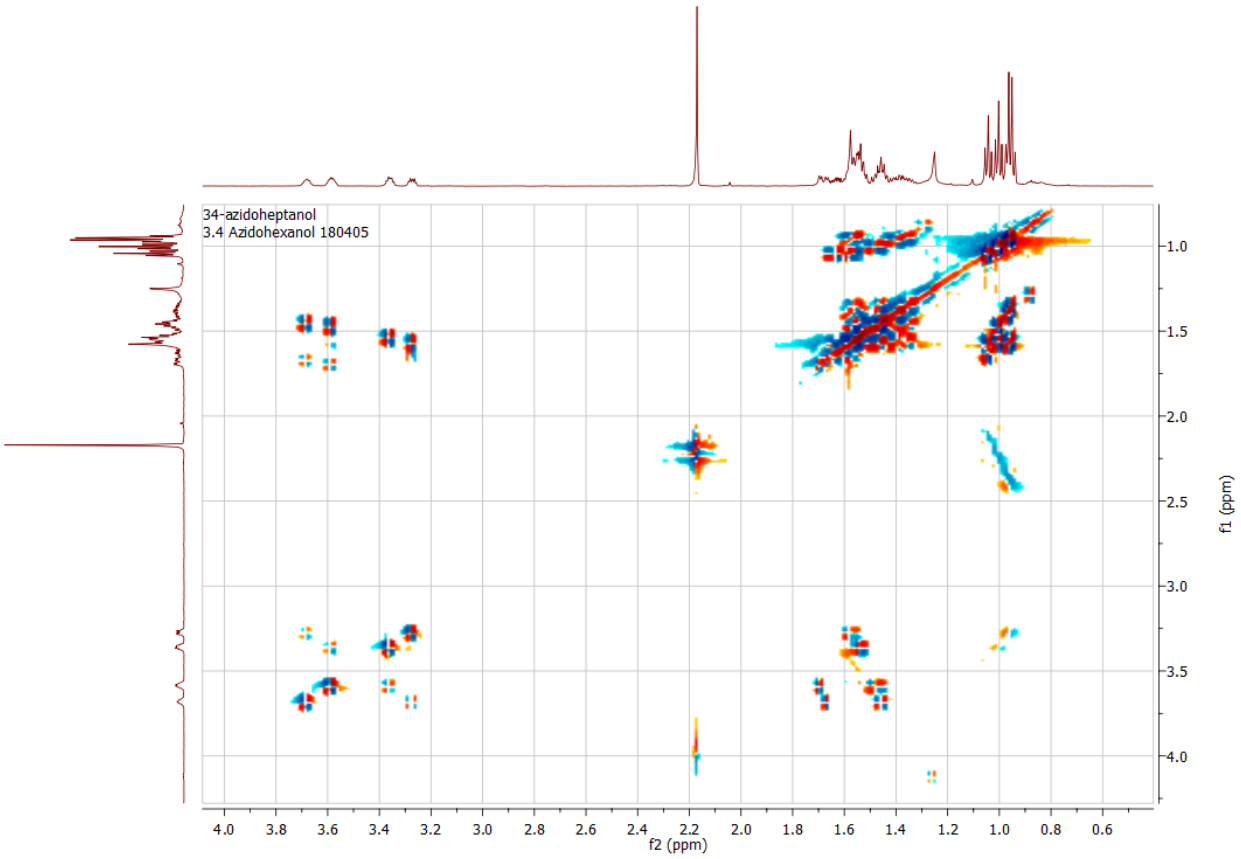


Supporting figure 15. HSQC spectrum, applied for the identification of regioisomers of **33** (Spectrum analysed with TopSpin 4.0).

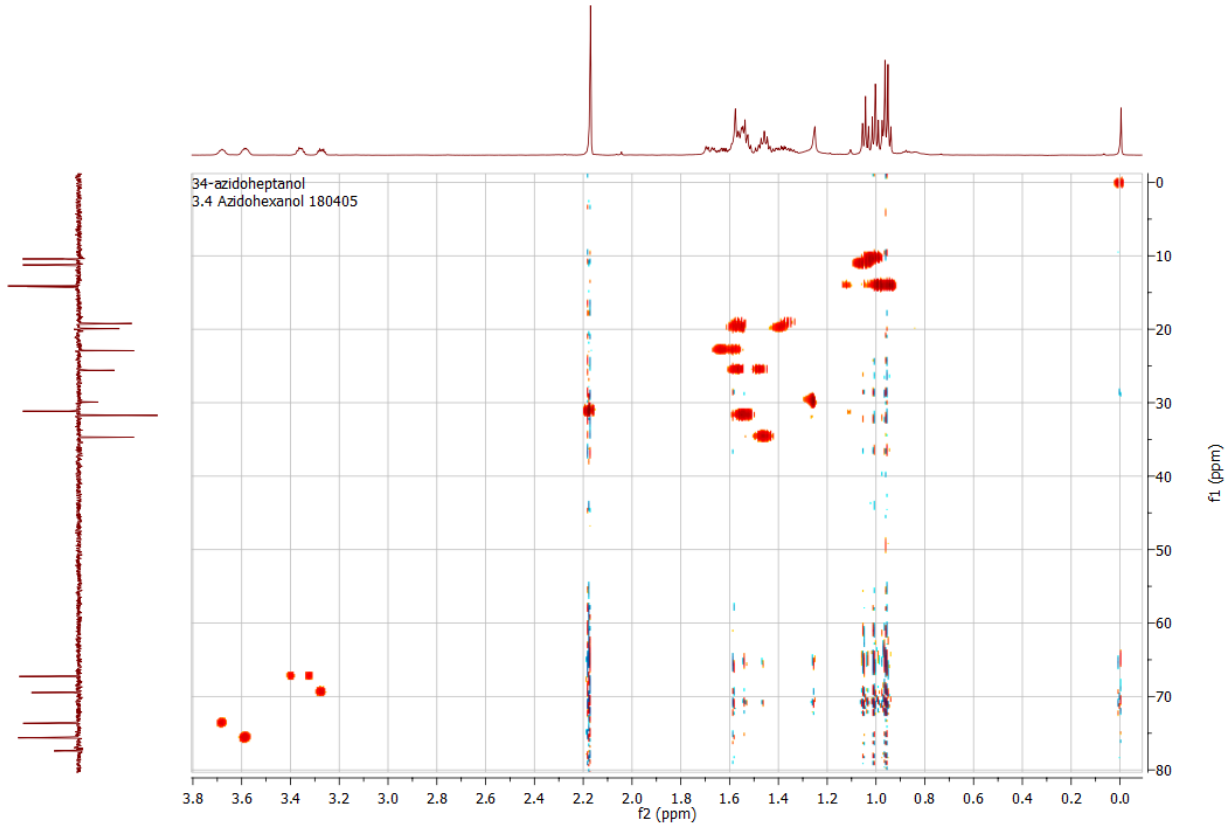
Supporting figure 16. ^{1}H NMR (600 MHz, $CDCl_3$) of the regioisomeric mixture of azidoheptanol **34** (Spectrum analysed with MestReNova 6).



Supporting figure 17. DEPT (151 MHz, $CDCl_3$) of the regioisomeric mixture of azidoheptanol **34** (Spectrum analysed with MestReNova 6).

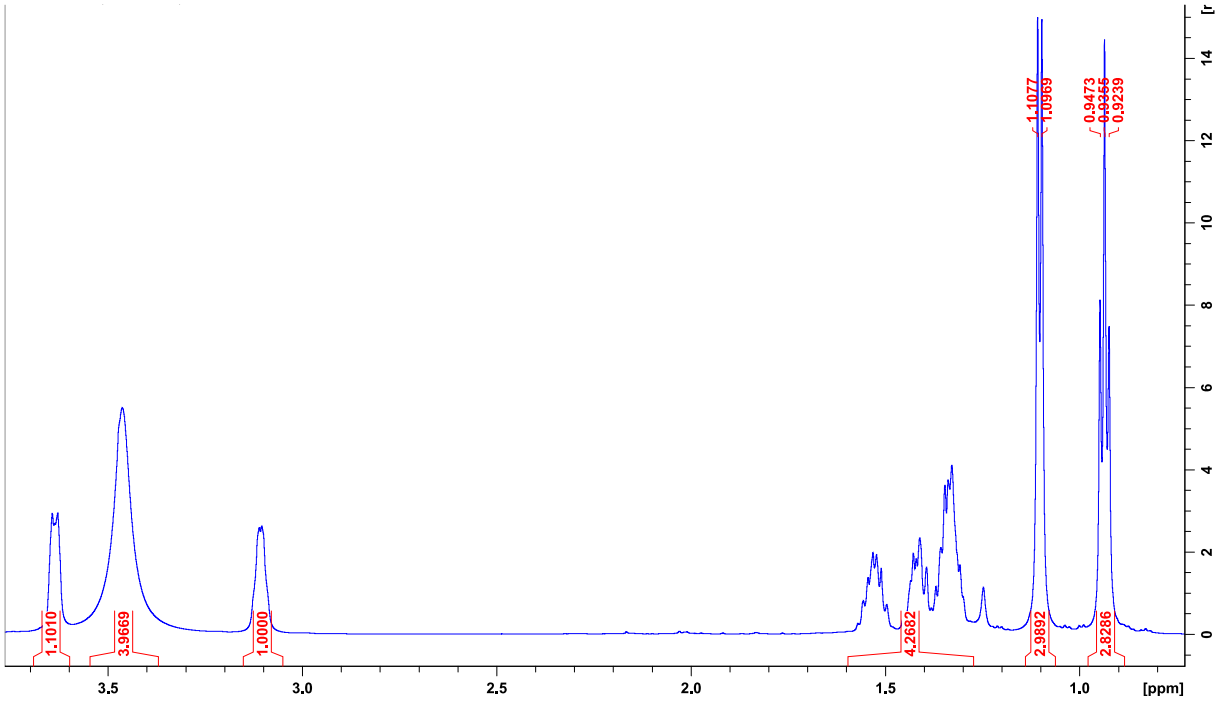


Supporting figure 18. COSY spectrum with double quantum filter, applied for the identification of regioisomers of **34** (Spectrum analysed with MestReNova 6).

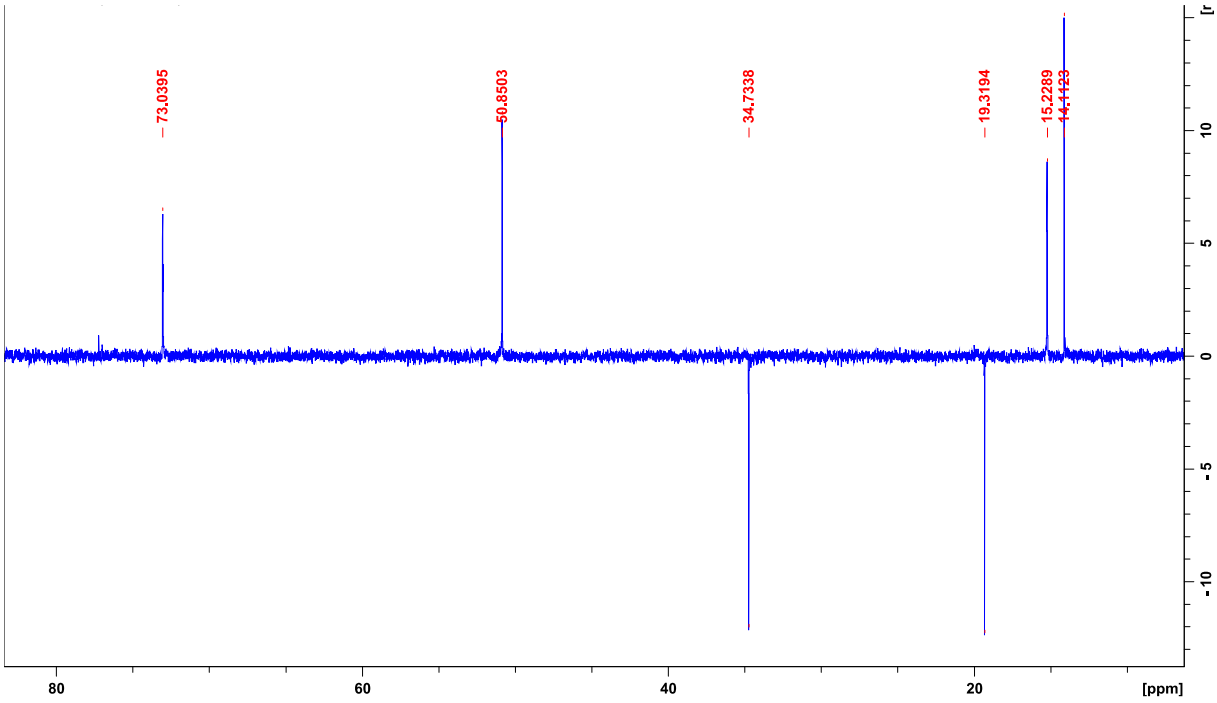


Supporting figure 19. HSQC spectrum, applied for the identification of regioisomers of **34** (Spectrum analysed with MestReNova 6).

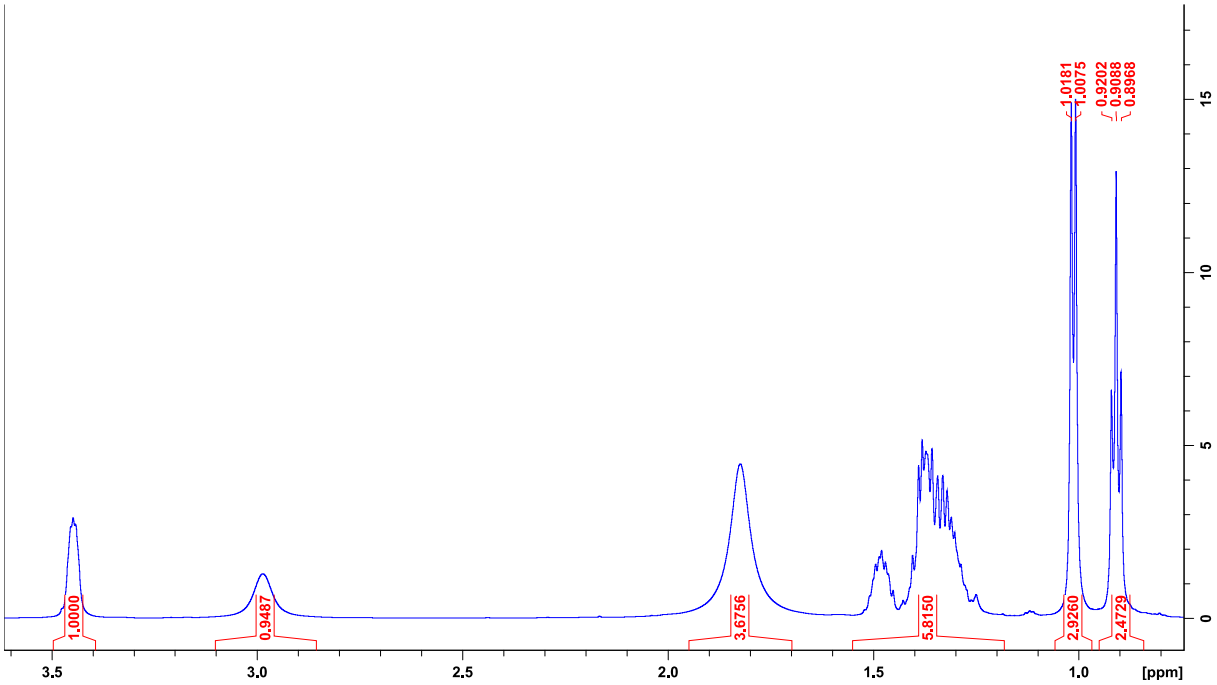
Amino alcohols



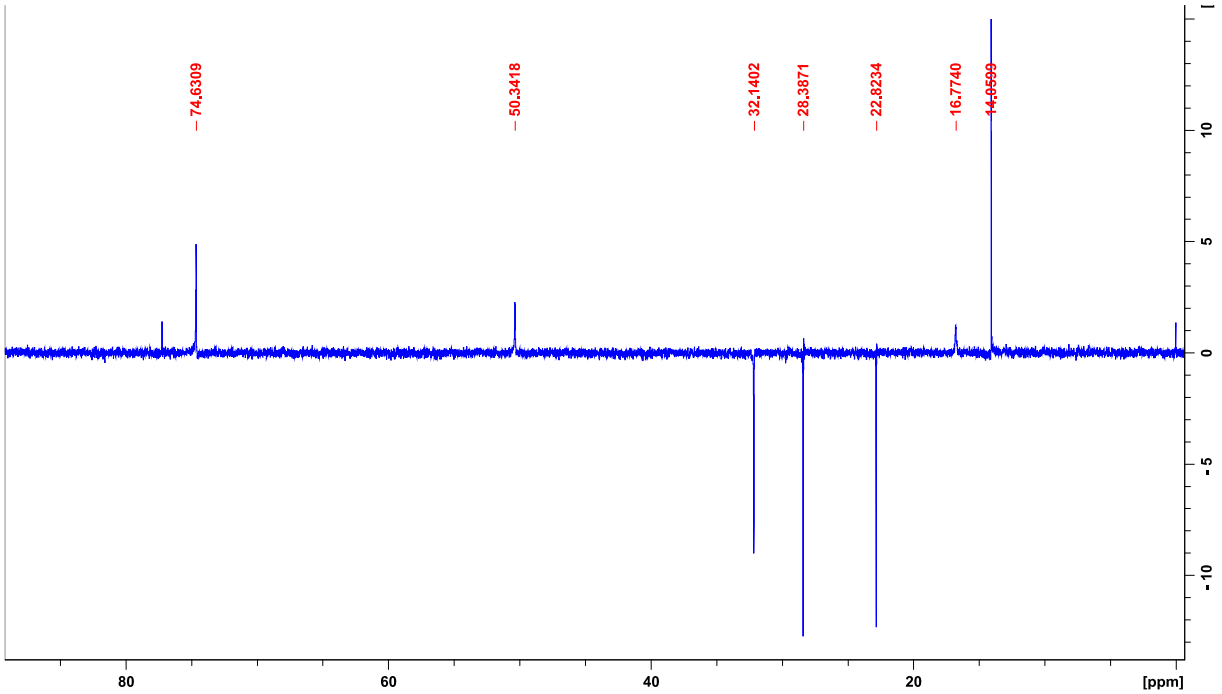
Supporting figure 20. ^{1}H NMR (600 MHz, $CDCl_3$) of 2-amino-3-hydroxyhexane (**41a**) (Spectrum analysed with TopSpin 4.0).



Supporting figure 21. DEPT (151 MHz, $CDCl_3$) of 2-amino-3-hydroxyhexane (**41a**) (Spectrum analysed with TopSpin 4.0).

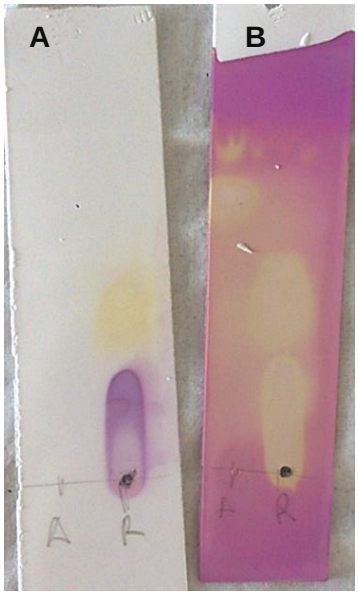


Supporting figure 22. ^1H NMR (600 MHz, $CDCl_3$) of 2-amino-3-hydroxyheptane (**42a**) (Spectrum analysed with TopSpin 4.0).



Supporting figure 23. DEPT (151 MHz, $CDCl_3$) of 2-amino-3-hydroxyheptane (**42a**) (Spectra analysed with TopSpin 4.0).

Thin Layer Chromatography

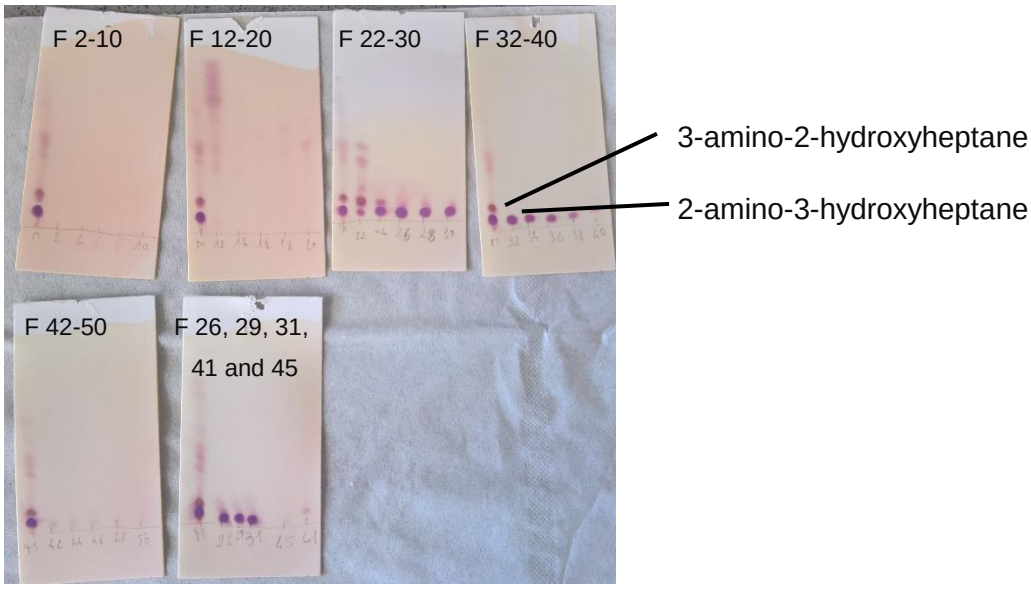


2-azido-3-hydroxyheptane

Hydrazine

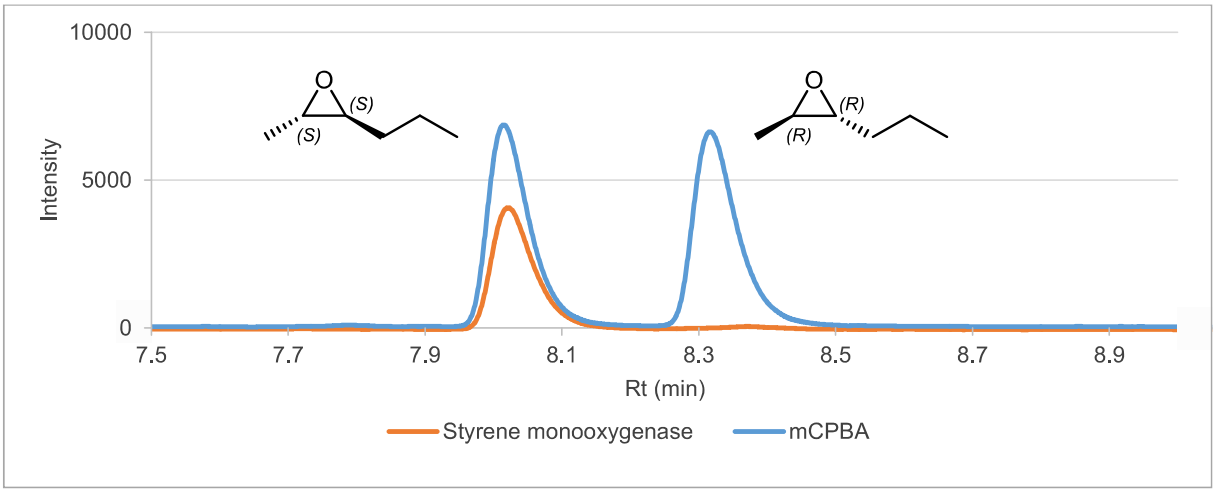
2-amino-3-hydroxyheptane

Supporting figure 24. Exemplary TLCs where the specific stain for amino groups (ninhydrin, A) and the general potassium permanganate stain (B) were used to follow the hydrogenation reaction. In B 2-azido-3-hydroxyheptane is not visible anymore in TLC.

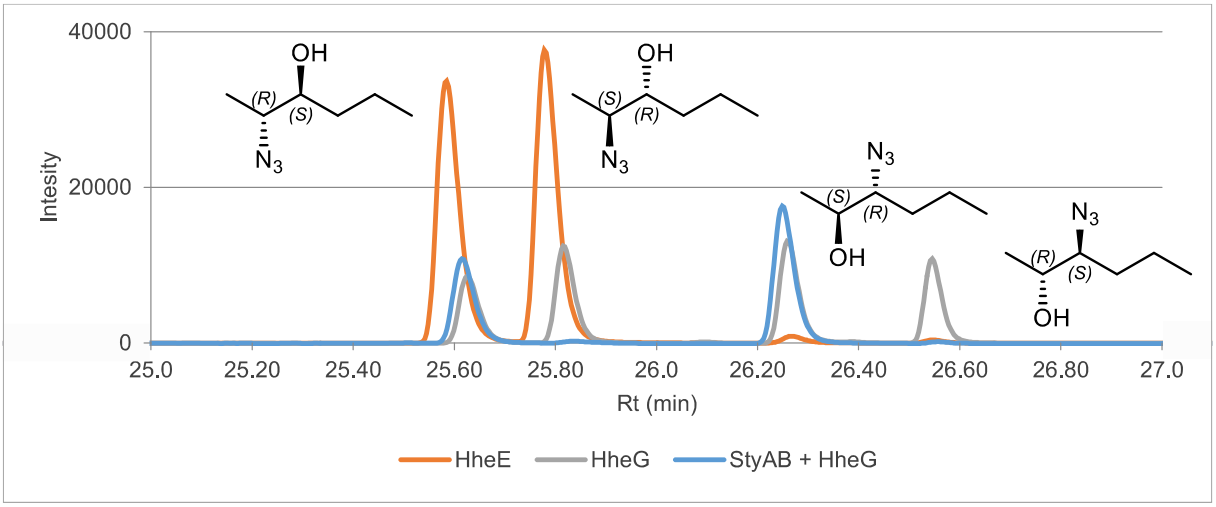


Supporting figure 25. Exemplary column chromatography for the purification of 2-amino-3-hydroxyheptane. Most impurities eluted in the first fractions (F 1-22).

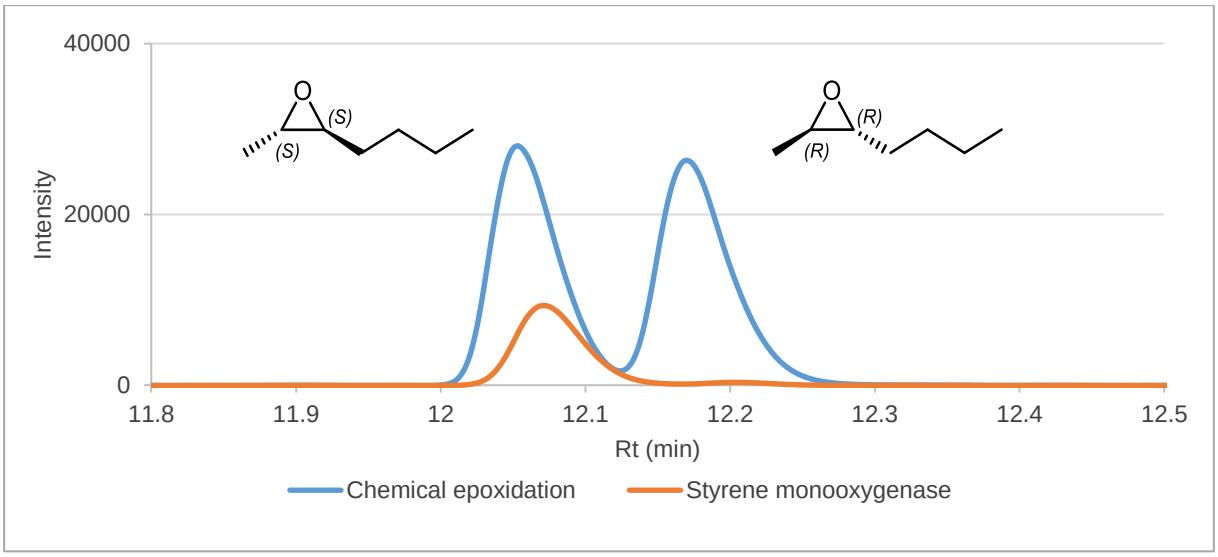
Chiral GC chromatograms



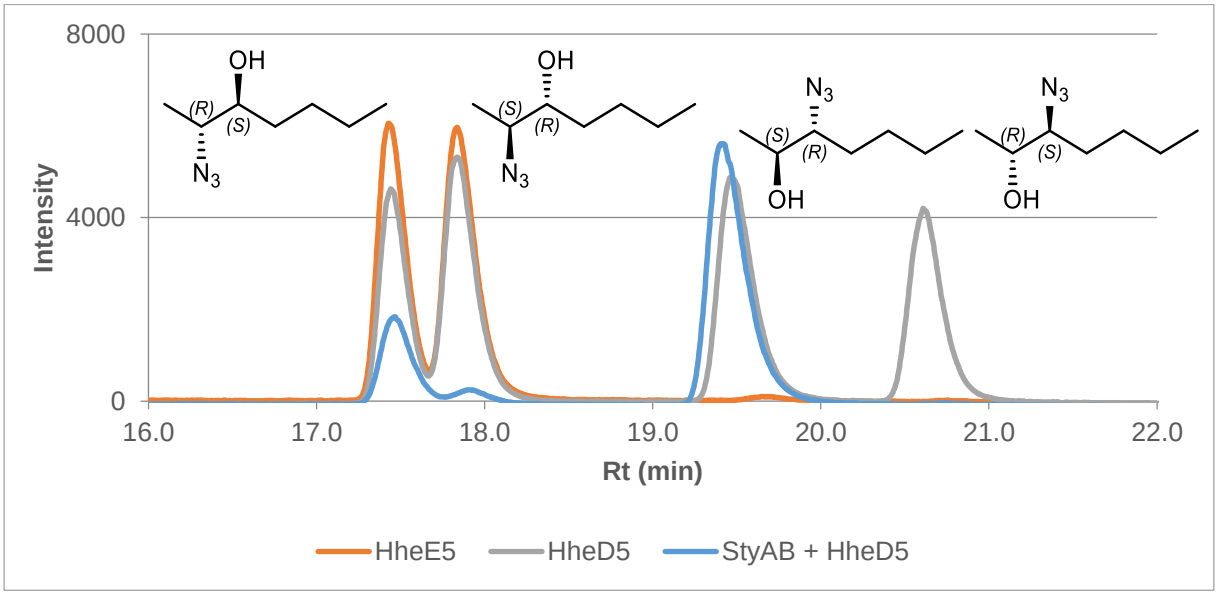
Supporting figure 26. Chromatograms of *trans*-2,3-epoxyhexane (**21**) synthesized by chemical epoxidation (blue) or by the styrene monooxygenase from *Rhodococcus* sp. ST-10 (orange).



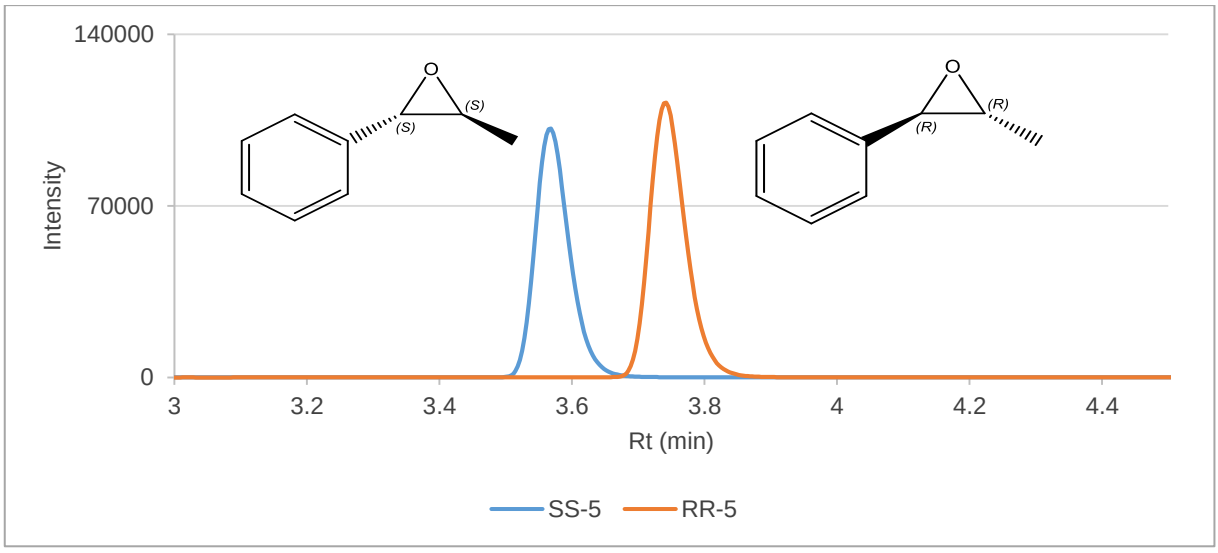
Supporting figure 27. Chromatograms of 2-azido-hexan-3-ol (**31a**) and 3-azido-hexan-2-ol (**31b**) enantiomers synthesized by a regioselective HHDH (orange), by a non-regioselective HHDH (grey) and by a non-regioselective HHDH converting (*S*,*S*)-**21** (blue).



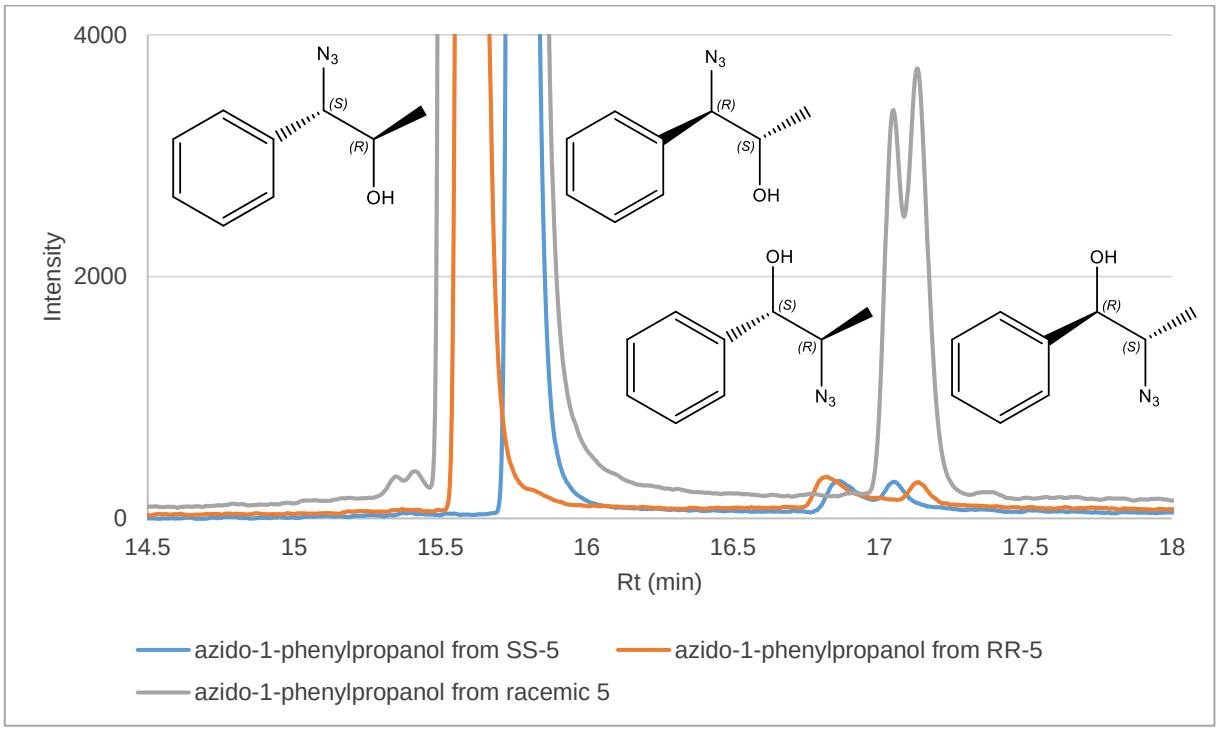
Supporting figure 28. Chromatograms of *trans*-2,3-epoxyheptane (**22**) synthesized by chemical epoxidation (blue) or by the styrene monooxygenase from *Rhodococcus* sp. ST-10 (orange).



Supporting figure 29. Chromatograms of 2-azido-heptan-3-ol (**32a**) and 3-azido-heptan-2-ol (**32b**) enantiomers synthesized by a regioselective HHDH (orange), by a non-regioselective HHDH (grey) and by a non-regioselective HHDH converting (*S,S)*-**22** (blue)**.**

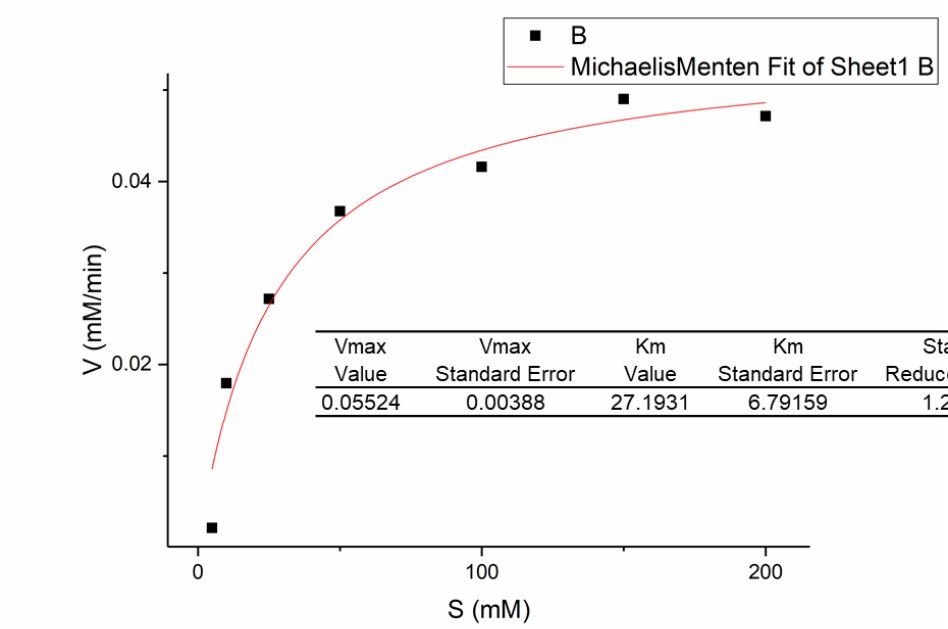


Supporting figure 30. Chromatograms of the commercially available pure enantiomers of *trans*-β-methyl styrene oxide (**25**) [(*S,S*)-**25** in blue and (*R,R*)-**25** in orange]



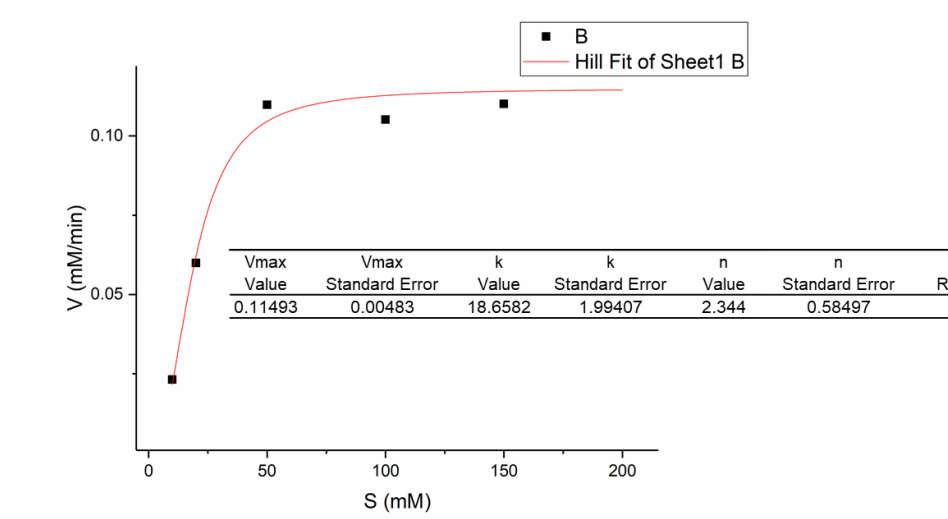
Supporting figure 31. Chromatograms of 2-azido-1-phenylpropan-1-ol (**35a**) and 1-azido-1-phenylpropan-2-ol (**35b**) enantiomers synthesised using the commercially available (*S,S*)-**25** (blue), (*R,R*)-**25** (orange) and the racemic mixture of **25** (grey) as substrates.

Enzyme Kinetics



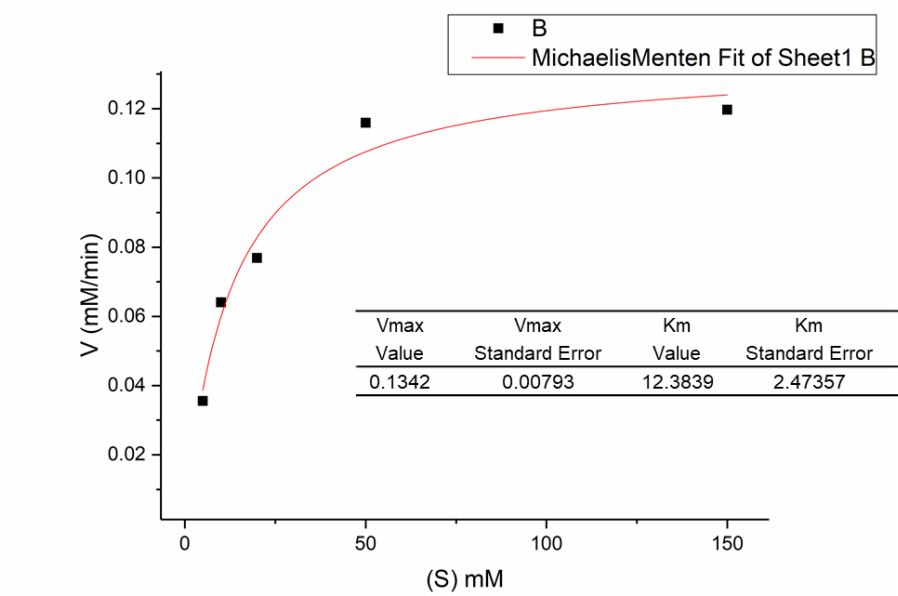
Vmax Value	Vmax Standard Error	Km Value	Km Standard Error	Statistics Reduced Chi-Sqr	Statistics Adj. R-Square
0.05524	0.00388	27.1931	6.79159	1.27E-05	0.95644

Supporting figure 32. HheE Michaelis-Menten kinetics for *trans*-2,3-epoxyhexane (**21**) and 50 mM of NaN_3



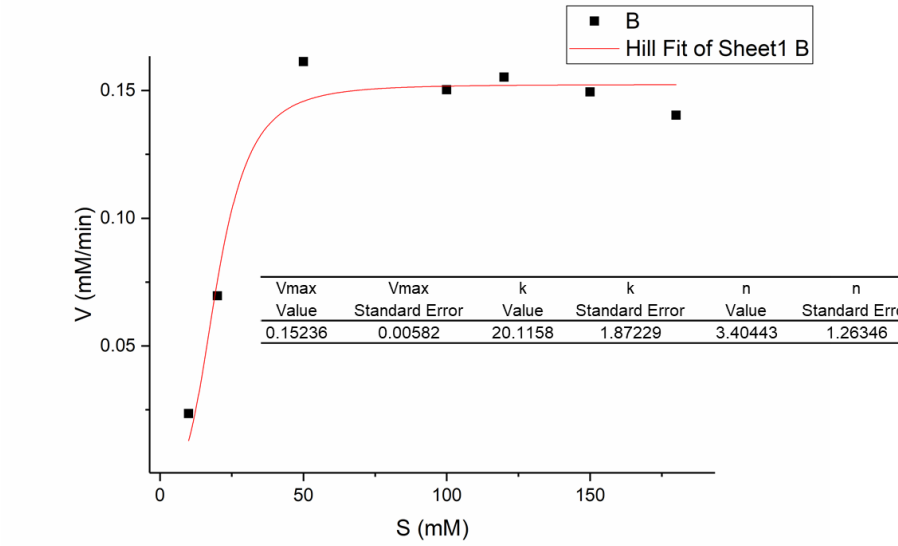
Vmax Value	Vmax Standard Error	k Value	k Standard Error	n Value	n Standard Error	Statistics Reduced Chi-Sqr	Statistics Adj. R-Square
0.11493	0.00483	18.6582	1.99407	2.344	0.58497	5.50E-05	0.96282

Supporting figure 33. HheE Hill kinetics for *trans*-2,3-epoxyheptane (**22**) and 50 mM of NaN_3.



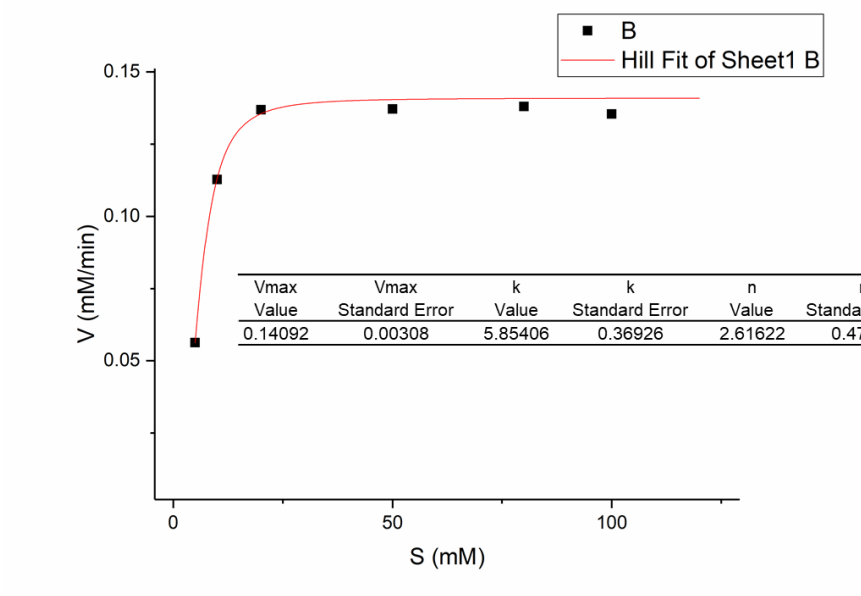
Vmax Value	Vmax Standard Error	Km Value	Km Standard Error	Statistics Reduced Chi-Sqr	Statistics Adj. R-Square
0.1342	0.00793	12.3839	2.47357	5.01E-05	0.96054

Supporting figure 34. HheE Michaelis-Menten kinetics for NaN_3 and 50 mM of *trans*-2,3-epoxyhexane (**21**).



Vmax Value	Vmax Standard Error	k Value	k Standard Error	n Value	n Standard Error	Statistics Reduced Chi-Sqr	Statistics Adj. R-Square
0.15236	0.00582	20.1158	1.87229	3.40443	1.26346	1.38E-04	0.9514

Supporting figure 35. HheE5 Hill kinetics for *trans*-2,3-epoxyheptane (**22**) and 50 mM of NaN_3.



Vmax Value	Vmax Standard Error	k Value	k Standard Error	n Value	n Standard Error	Statistics Reduced Chi-Sqr	Statistics Adj. R-Square
0.14092	0.00308	5.85406	0.36926	2.61622	0.47494	3.88E-05	0.96205

Supporting figure 36. HheE5 Hill kinetics for NaN_3 and 50 mM of *trans*-2,3-epoxyheptane (**22**).

ACKNOWLEDGEMENTS

This project has received funding from the European Union's Horizon 2020 MSCA ITN-EID program under grant agreement No 634200 and it was in collaboration with TU Delft and Enzymicals.

I'd like to give the first and special thanks to Prof. Dr. Anett Schallmey for allowing me into her laboratory even though I had a background from a very different topic. For teaching me the fundamentals of biotechnology with her always smart, useful and practical advices that guided me throughout the whole PhD work.

I would like to thank Prof. Dr.-Ing. Antje Spiess for agreeing to review my thesis and for being a member of the PhD evaluation committee.

I would also like to thank Prof Dr. Dieter Jahn for being the third PhD committee member.

I would like to thank Frank Hollmann for the supervision during my external stay at TU Delft.

A big thank to Enzymicals for hosting the second half of my PhD and in particular to Dr. Philipp Süss and Dr. Melinda Fekete as they were always ready to answer my 999 daily questions.

I would like to thank the amazing ESRs of the Biocascades project: Aline, Elisa, Federica, Florian, Gaurav, Juraj, Laura, Lisa, Olha and Raquel for the nice time together and emotional support. Furthermore, all the people involved in the project for the very interesting scientific discussion and "tips and tricks".

I want to thank all the people I met in the different research group for always including me like part of the family:

TU Braunschweig: Willem, Patrick, Kathi, Marcus, Jenny, Jana, Hauke, Jon Alex, Rita, Steffi Constantin and especially Julia (and Emmi but only when asleep) and Janine for the contributions to this thesis and scientific publications.

TU Delft: Sébastien, Marine, Morten, Georg, Tiago, Hanna, Florian, Caroline, Sandy, Lìa, Remco and Marten.

Ezymicals and Uni Greifswald: Philipp, Melinda, Rainer, Henrike, Sabine, Eric, Falko, Friedemann, Dennis, Sebastian, Sven, Thobias, Ute,Thomas, Lisa, Vishnu, Lukas, Katja, Miriam, Chris, Emil, Ayad, Henrik, Eva, Sasha, Moritz and Ina.

I will probably forget someone but I hope you would excuse me, it was quite an adventure.

Thanks Emilia for resisting and still being around even if you met me in the last agitated months of my PhD.

A special gratefulness to thesaurus.com for subsidizing this written dissertation by furnishing exotic english words of all sorts (promptly extirpated by Julia and Anett).

Thanks to my friends for being always there every time I went back to visit my hometown to disrupt their routines.

Finally, a huge thanks to my family always present and supporting the “whatevers ciencestuff” I do.
Come in ogni tesi un ringraziamento speciale ad Andrea e Patrizia per credere in me anche se ancora non s’è capito che sto a fa. Un grazie alle nonne Anna e Nella per il supporto e per le doti culinarie sempre di alto livello. E un grazie alla sorellina che ormai è una dottorona pure lei e quello che sto a fa lo capisce un pelo meglio.

For the reader that made it this far you probably found some mistakes: typos or experiments that could be done differently or maybe you are another PhD student trying to repeat some of the work done. Never give up, because without mistakes there is no science. And if my word is not enough, after all I am also a human that makes mistakes, I will add some quotes from very successful people which by chance are also human that make mistakes.

- "Negative results are just what I want. They’re just as valuable to me as positive results. I can never find the thing that does the job best until I find the ones that don’t." **(Thomas A. Edison)**
- "Results! Because, man, I have gotten a lot of results. I know several thousand things that won`t work." **(Thomas A. Edison)**
- "Anyone who has never made a mistake has never tried anything new." **(Albert Einstein)**
- "Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth." **(Jules Verne)**
- "An expert is a man who has made all the mistakes which can be made, in a narrow field." **(Niels Bohr)**
- "Every great improvement has come after ten repeated failures. Virtually nothing comes right in the first time." **(Charles F. Kettering)**
- "Successes teach you nothing. Failures teach you everything. Making mistakes is the most important thing you can do." **(James Dyson)**

Academic CV

Personal information

Name:	Elia Calderini
Birthday:	20. August 1990
Place of birth:	Città di Castello, Italy
Nationality:	Italian

Employment history

Postdoctoral researcher – Graz University of Technology (2019 – present)

PhD Biotechnology – Technical University of Braunschweig (2015 – tentative date: Sept 2019)

Thesis title: A chemo-enzymatic cascade for the production of enantiopure aminoalcohols.

Education history

MSc Industrial Biotechnology – University of Turin (2012 – 2014)
Master thesis (done at Technical University of Denmark): Proteomic analysis of oxidative stress response in lactic acid bacteria.
Mark: 110/110 *cum laude*

BSc Biotechnology – University of Perugia (2009 – 2012)
Bachelor thesis: “Comparative gene expression profiling between wild-type and GILZ knock out spermatogonia in C57 BL6 mouse”.
Mark: 110/110 *cum laude*

Peer-reviewed publications:

Calderini E., Wessel J., Süss P., Schrepfer P., Wardenga R., Schallmey, A. Selective Ring-Opening of Di-Substituted Epoxides Catalysed by Halohydrin Dehalogenases. *ChemCatChem* **2019**, *11* (8), 2099–2106. https://doi.org/10.1002/cctc.201900103.

Calderini E., Celebioglu H.U., Villarroel J., Jacobsen S., Svensson B., Pessione E., Comparative proteomics of oxidative stress response of *Lactobacillus acidophilus* NCFM reveals effects on DNA repair and cysteine de novo synthesis. PROTEOMICS **2017**, 1600178. https://doi.org/10.1002/pmic.201600178.

Additional research achievements:

Oral presentation: Elia Calderini, Julia Koopmeiners, Rainer Wardenga and Anett Schallmey. Selective ring opening of di-substituted epoxides catalysed by halohydrin dehalogenases. 2017, **Microbiology and Infection 2017** - 5th Joint Conference of the DGHM & VAAM, Würzburg, Germany.

Oral presentation: Elia Calderini; Philipp Süss; Rainer Wardenga and Anett Schallmey. A chemo-enzymatic cascade for the production of enantiopure amino alcohols. **2018, 8th International Cebitec Research Conference**, Bielefeld, Germany.